

# A REDSHIFT SURVEY OF NEARBY GALAXY GROUPS: THE SHAPE OF THE MASS DENSITY PROFILE

ANDISHEH MAHDVI<sup>1</sup>

Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822

AND

MARGARET J. GELLER<sup>2</sup>

Smithsonian Astrophysical Observatory, 60 Garden St., Cambridge, MA 02138

*Accepted for publication in the Astrophysical Journal*

## ABSTRACT

We constrain the mass profile and orbital structure of nearby groups and clusters of galaxies. Our method yields the joint probability distribution of the density slope  $n$ , the velocity anisotropy  $\beta$ , and the turnover radius  $r_0$  for these systems. The measurement technique does not use results from N-body simulations as priors. We incorporate 2419 new redshifts (included here) in the fields of 41 systems of galaxies with  $z < 0.04$ . The new groups have median velocity dispersion  $\sigma = 360 \text{ km s}^{-1}$ . We also use 851 archived redshifts in the fields of 8 nearly relaxed clusters with  $z < 0.1$ . Within  $R \lesssim 2r_{200}$ , the data are consistent with a single power law matter density distribution with slope  $n = 1.8\text{--}2.2$  for systems with  $\sigma < 470 \text{ km s}^{-1}$  and  $n = 1.6\text{--}2.0$  for those with  $\sigma > 470$  (95% confidence). We show that a simple, scale-free phase space distribution function  $f(E, L^2) \propto (-E)^{\alpha-1/2} L^{-2\beta}$  is consistent with the data as long as the matter density has a cusp. Using this DF, matter density profiles with constant density cores ( $n = 0$ ) are ruled out with better than 99.7% confidence.

## 1. INTRODUCTION

The galaxies in a cluster account for less than 10% of its total mass. X-ray observations of the hot intracluster gas support the notion that most of the matter is distributed in a smooth, dark halo. It is of clear interest to use observations to deduce not only the total mass, but also the structure of this halo. For example, the logarithmic derivative of the density can be compared with N-body simulations of structure formation, allowing constraints on dark matter physics. These simulations show that the inner regions of halos composed of collisionless dark matter have power law density profiles  $\rho \propto r^{-n}$  with  $n \sim 1 - 1.5$  (Navarro et al. 1997, NFW; Fukushige & Makino 1997; Nakano & Makino 1999; Moore et al. 1999). However, if the dark matter particles interact through the weak or the strong force, constant density cores develop within a fraction of a Hubble time (Burkert 2000; Davé et al. 2001; Balberg et al. 2002). Measurements of the slope  $n$  in groups and clusters of galaxies are thus a potential test of the collisionless nature of dark matter.

The methods for determining the masses of clusters fall into three broad categories: fluid dynamics, lensing, and stellar dynamics. The fluid dynamical methods, which rely on the X-ray properties of the intracluster medium, presuppose that the gas is either in hydrostatic equilibrium or in a time-independent cooling flow. Because measuring the gas temperature involves complex modeling of X-ray spectra, these methods offer at best  $\sim 50$  resolution elements across the brightest X-ray sources. Furthermore, recent X-ray observations indicate that the central regions of cooling flow clusters cannot be described self-consistently

by the standard models, leaving the shape of the mass profile within  $\sim 200 \text{ kpc}$  of the cluster center uncertain<sup>3</sup> (Markevitch et al. 1999; Soker et al. 2001; Fabian et al. 2001).

Lensing models detect either strong or weak gravitational deflection of the light from distant sources. Strong lensing allows direct estimation of the surface mass density from the spectacular but rare giant arcs, e.g. CL0024+16 (Tyson et al. 1998). Weak lensing estimates depend on statistical reconstruction of the surface density from distortions in the shapes of field galaxies (Clowe et al. 2000; Sheldon et al. 2001). These techniques are attractive because they make no assumptions about the equilibrium state of the cluster. However, they could suffer contamination from the large scale structure surrounding the cluster (Metzler, White, & Loken 2001). This contamination may account for the disagreement between X-ray and lensing masses, and makes the determination of the true shape of the density profile difficult.

Stellar dynamical methods grow out of the century-old tradition, beginning with Jeans and Eddington, of modeling the structure of star clusters. Here, instead of addressing a self-gravitating system of stars, one regards galaxies as point masses adrift in a larger sea of dark matter. The fundamental dynamical problem is then to calculate the spherically symmetric gravitational potential that causes the observed motions of the galaxies. Obviously, the analogy to stellar systems is far from perfect; the ratio of the size of a galaxy to that of its host cluster is typically  $\approx 15 \text{ kpc} / 1.5 \text{ Mpc} = 0.01$ , whereas for stars in a galaxy it is  $\approx 10^{-12}$ . As a result, interaction cross sections are much larger in galaxy clusters. Furthermore, although

<sup>1</sup> amahdavi@ifa.hawaii.edu

<sup>2</sup> mgeller@cfa.harvard.edu

<sup>3</sup> Throughout this paper we use  $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

stellar systems contain as many as  $10^{12}$  members, clusters rarely have more than  $\approx 500$  luminous members, and the most abundant systems—groups of galaxies—have closer to  $\approx 30$  (Carlberg et al. 1996; Mahdavi et al. 1999).

Still, if correctly applied, stellar dynamics can trace the gravitational potential of clusters at the largest and smallest scales as well as the other available techniques. Modeling of spherical infall patterns (Geller et al. 1999; Rines et al. 2000) can map the mass profile outside  $\approx 5$  Mpc. The equilibrium models we consider are sensitive to the shape of the dark matter halo in the innermost regions of clusters and groups.

The most popular equilibrium techniques make use of the moments of the data to constrain the depth or shape of the cluster potential. For example, the virial theorem yields an estimate of the total mass of the cluster (Heisler et al. 1985; Biviano et al. 1993; Oegerle et al. 1995; Carlberg et al. 1996; Girardi & Giuricin 2000):

$$M = \frac{3\pi}{2G} \sum_{i=1}^N v_{z,i}^2 R_h, \quad (1)$$

$$R_h = \sum_{i=1}^N R_i^{-1} \quad (2)$$

$$v_{z,i} = \frac{cz_i - \sum_{i=1}^N cz_i/N}{1 + \sum_{i=1}^N z_i}, \quad (3)$$

where  $v_z$  is the line-of-sight velocity in the center of mass frame (Danese et al. 1980),  $z_i$  is the redshift of the  $i$ th galaxy,  $R_i$  is its projected distance from the cluster center,  $c$  is the speed of light, and  $G$  is Newton's constant.

The velocity moments may also be used to infer the structure of the potential in greater detail, using either analytic galaxy distribution functions (Kent & Gunn 1982) or the Jeans equation for a collisionless stellar system (Fabricant et al. 1989; Binney & Tremaine 1987; den Hartog & Katgert 1996; Carlberg et al. 1997),

$$\frac{1}{\nu} \frac{d(\nu \sigma_r^2)}{dr} + \frac{2(\sigma_r^2 - \sigma_t^2)}{r} = -\frac{d\Phi}{dr}, \quad (4)$$

where  $r$  is the true three-dimensional distance from the cluster center,  $\nu(r)$  is the number distribution of the galaxies,  $\sigma_t(r)$  and  $\sigma_r(r)$  are the tangential and the radial velocity dispersions, and  $\Phi$  is the gravitational potential. To interpret the data, one usually chooses a form for  $\Phi(r)$  and the anisotropy parameter  $\beta(r) \equiv 1 - \sigma_t^2/\sigma_r^2$ , and projects the solution to the Jeans equation to obtain  $\sigma_z(R)$ , the theoretical line-of-sight velocity dispersion profile. This profile may be compared with a real cluster by splitting the galaxies into radial bins and calculating  $\sigma_z^2 \propto \sum v_z^2$  in each bin.

There are several disadvantages to using the velocity moments to constrain  $\Phi$ . First, the quality of the observed velocity dispersion profile  $\sigma_z(R)$  is usually poor for clusters. Even with several hundred velocities, dividing the galaxies into radial bins produces a noisy profile that is not very informative about the radial variation of  $\sigma_z$ . Second, even in the ideal limit of a perfectly observed  $\sigma_z(R)$ , vastly different combinations of  $\Phi(r)$  and  $\beta$  can yield similar solutions to equation (4) (Binney & Tremaine 1987). Third, even if a unique solution is possible (e.g., by assuming a constant mass-to-galaxy ratio,  $\nabla^2\Phi \propto \nu$ ), there is no

guarantee that the solutions satisfy the requirement that the phase space density of the member galaxies be everywhere positive or zero (van der Marel et al. 2000). Finally, it is desirable to avoid binning the velocities altogether in order to make as powerful a use of the scarce data as possible. We therefore turn to maximum likelihood methods, which avoid some of the problems of the Jeans equations as applied to discrete systems (Merritt & Saha 1993; van der Marel et al. 2000). Constructing and maximizing a suitable likelihood function guarantees a positive definite galaxy phase space distribution.

Here we apply the maximum likelihood method to nearby ( $z \lesssim 0.1$ ) systems of galaxies. Our goal is to derive joint constraints on  $\Phi$  and  $\beta$  for poor groups of galaxies as well as for rich clusters. To this end we conduct deep optical observations of the RASSCALS X-ray emitting galaxy groups (Mahdavi et al. 2000). We also assemble a catalog of published redshifts in 8 nearby relaxed clusters of galaxies (§2). Using these samples, we construct an ensemble group and ensemble cluster which serve to constrain the galaxy phase space distribution in each type of system (§3). This distribution then serves as a maximum likelihood estimator of the gravitational potential and of the orbital structure of the galaxy population. In particular, we calculate joint five-dimensional confidence volumes in the central anisotropy, central matter density slope, total mass, matter core radius, and interloper fraction (§4). We discuss the implications of our work (§5) and conclude (§6).

## 2. DATA

Here we describe the data acquisition. First, we discuss our observations of primarily poor groups of galaxies. From these data we obtain a statistically complete sample of 2419 optical spectra in the fields of 41 systems. We also assemble a redshift catalog of galaxies in relaxed rich clusters from the literature.

### 2.1. New Observations

Our redshift measurement targets (Table 1) are drawn from Mahdavi et al. (2000), who cross-correlate the Center for Astrophysics Redshift Survey with the ROSAT All-Sky Survey (RASS) to construct the RASSCALS catalog. The RASSCALS are a statistically complete, redshift-selected sample of galaxy groups with at least 5 members brighter than a Zwicky blue magnitude  $m_Z = 15.5$ , comparable to a red magnitude  $m_R \approx 14.4$ . The groups lie within a redshift range  $3000 \leq cz \leq 12000$  km s $^{-1}$ , and have either X-ray luminosities or upper limits derived from the RASS. The diffuse X-ray emission coincident with the detected groups suggests that they are physically bound systems of galaxies.

For this study, we select all X-ray emitting RASSCALS with right ascension  $8 < \alpha_{2000} < 16$ , or  $\alpha_{2000} > 22$  and  $\alpha_{2000} < 2$ , and declination  $\delta_{2000} > -8$ . We restrict ourselves to systems well away from the local supercluster,  $cz > 5100$  km s $^{-1}$ , and attempt to ensure that they are group-like by limiting ourselves to X-ray luminosities  $10^{42} < L_X < 10^{43} h_{100}^{-2}$  erg s $^{-1}$ . We also omit the Hickson (1982) compact groups 51 and 58. In addition to this 93% complete sample of 30 X-ray emitting groups, we randomly select 11 groups without RASS X-ray emission, all

located within the same  $(\alpha_{2000}, \delta_{2000}, z)$  bounds. There are 41 systems in all.

To constrain the variation of the velocity dispersion, and hence the mass, with radius, it is necessary to identify group members to a large distance from the center of the gravitational potential. We therefore identify galaxies brighter than  $m_R \approx 15.4$  within 1.5 Mpc of the center of all 41 groups as spectroscopic targets. We include all of these galaxies in our spectroscopic survey. Occasionally, we do not reobserve galaxies with a known redshift  $z$  such that  $|cz - c\bar{z}| > 4000 \text{ km s}^{-1}$ , where  $\bar{z}$  is the redshift of the group as listed in Mahdavi et al. (2000). In other words, we sometimes ignore spectroscopically confirmed foreground and background galaxies.

The total area of the sky covered by our spectroscopic survey is  $\approx 184$  square degrees. Thus the photometric calibration of plate-derived magnitudes is difficult, and the magnitude limit is only an estimate from the Palomar Observatory Sky Survey (POSS) photographic plates, with a typical scatter of 0.5 mag from system to system. Because we view the galaxies as test particles embedded in a gravitational potential, completeness is not as important as it would be if we wanted to measure the group luminosity function. Nevertheless, our sampling uses a strict magnitude cutoff in the field of each group, with no biases against galaxies close together on the sky.

We obtained spectra of the 2419 target galaxies with the FAST single object spectrograph (Fabricant et al. 1998) on the 1.5m Tillinghast Reflector at Whipple Observatory on Mount Hopkins, AZ. Using a 300 line  $\text{mm}^{-1}$  grating, a  $3''$  slit, and a spectral coverage of  $3940\text{\AA}$  centered at  $5500\text{\AA}$ , we had a typical resolution of  $4\text{\AA}$ . A data reduction pipeline incorporating the cross-correlation of the spectra with rest-frame galaxy templates (Tonry & Davis 1979; Kurtz et al. 1992; Kurtz & Mink 1998) gave redshifts with an average uncertainty of  $\pm 40 \text{ km s}^{-1}$ . Roughly 20% of these data were reported in Mahdavi et al. (1999); for completeness, we list all 2419 redshifts in Table 2.

## 2.2. Clusters

To compare the mass distribution of poor systems of galaxies and nearby rich clusters, we assemble a catalog of clusters close to dynamical equilibrium. This requirement is important, because the techniques we use below (and in fact all other methods relying on equilibrium dynamics) assume that the phase space distribution of the member galaxies is constant with time. Of course, at some level all clusters have residual substructure; our best hope is to examine those systems with the smallest departures from equilibrium.

It is surprising just how rare equilibrium systems are. For example, of the 10 brightest objects in the ROSAT Bright Cluster Sample (Ebeling et al. 1998, BCS), only two clusters, Abell 1795 and Abell 2029, make it to our list in Table 3. Our criteria are not stringent: we require that published studies of the X-ray emitting gas and the cluster galaxy population lack clear evidence of ongoing mergers. For example a strong bimodality in the X-ray or optical light distribution is a signal that the system is not close to dynamical equilibrium. We also require a large number ( $> 70$ ) of measured velocities in the field of each cluster; we search the NASA Extragalactic Database (NED) for

published redshifts within 1.5 Mpc of the cluster center.

The dynamical state of some of the clusters in our list is ambiguous. For example, the central dominant galaxy in Abell 2029 has a large peculiar velocity  $\sim 500 \text{ km s}^{-1}$  in the center of mass frame (Oegerle et al. 1995), but the cluster has a regular X-ray and optical morphology and shows no statistically significant substructure (Slezak et al. 1994; Oegerle et al. 1995). Many of the clusters we select contain cooling flows, which should arise only under equilibrium conditions (Allen 1998). However, not all cooling flow clusters have equilibrium galaxy populations; the most obvious example is Abell 85, which exhibits quite an irregular galaxy velocity structure (Durret et al. 1998). Abell 2199, another cooling flow cluster, is regular in both the X-ray and optical data within 1.5 Mpc. However, it is on course to collide with Abell 2197 and several other small galaxy groups located beyond 2 Mpc (Rines et al. 2001).

The clusters in our catalog have a mean redshift 0.04, and therefore complement the clusters of the Canadian Network for Observational Cosmology (Yee et al. 1996; Carlberg et al. 1996, 1997; van der Marel et al. 2000, CNOC), which are at  $z \approx 0.3$  on the average. The two samples differ in character, however: at least six of the CNOC clusters exhibit obvious substructure or asymmetry in X-ray images (Neumann & Bohringer 1997; Lewis et al. 1999), and possibly also in weak lensing maps (Clowe et al. 2000). The CNOC mass determinations rely on the assumption that equilibrium techniques average over the substructure without biasing the final result. To a lesser extent, this assumption underlies our study as well; the absence of substructure still does not guarantee that the galaxy phase space distribution has arrived at a steady state.

## 3. THE PHASE SPACE DIAGRAM

In this section we isolate the group and cluster members from foreground and background galaxies. To conduct the dynamical modeling, we combine the systems into two ensembles.

### 3.1. Membership

A redshift survey in the field of a particular cluster is always contaminated by foreground and background galaxies. Some outliers in velocity space are easily identifiable. Unrelated galaxies with velocities close to the cluster mean, however, are not as cleanly separated. Nor is it possible to identify a redshift range  $(z_1, z_2)$  which contains only the cluster members, because interlopers—unrelated galaxies with large peculiar motions—could well be projected into this redshift range.

We adopt a twofold approach to the membership problem. Here we exclude the most probable foreground and background galaxies by the method of Zabludoff, Huchra, & Geller (1990). Then in §4.4 we follow van der Marel et al. (2000) in folding the existence of the remaining interlopers directly into our dynamical model.

The Zabludoff et al. (1990) method consists entirely in making sure that in a sorted list of member velocities  $v_{z,i}$ , no two neighbors have  $\Delta v_z > \sigma_z$ , where  $\sigma_z$  is the line-of-sight velocity dispersion of the cluster. This approach is not the most sophisticated method available. There is

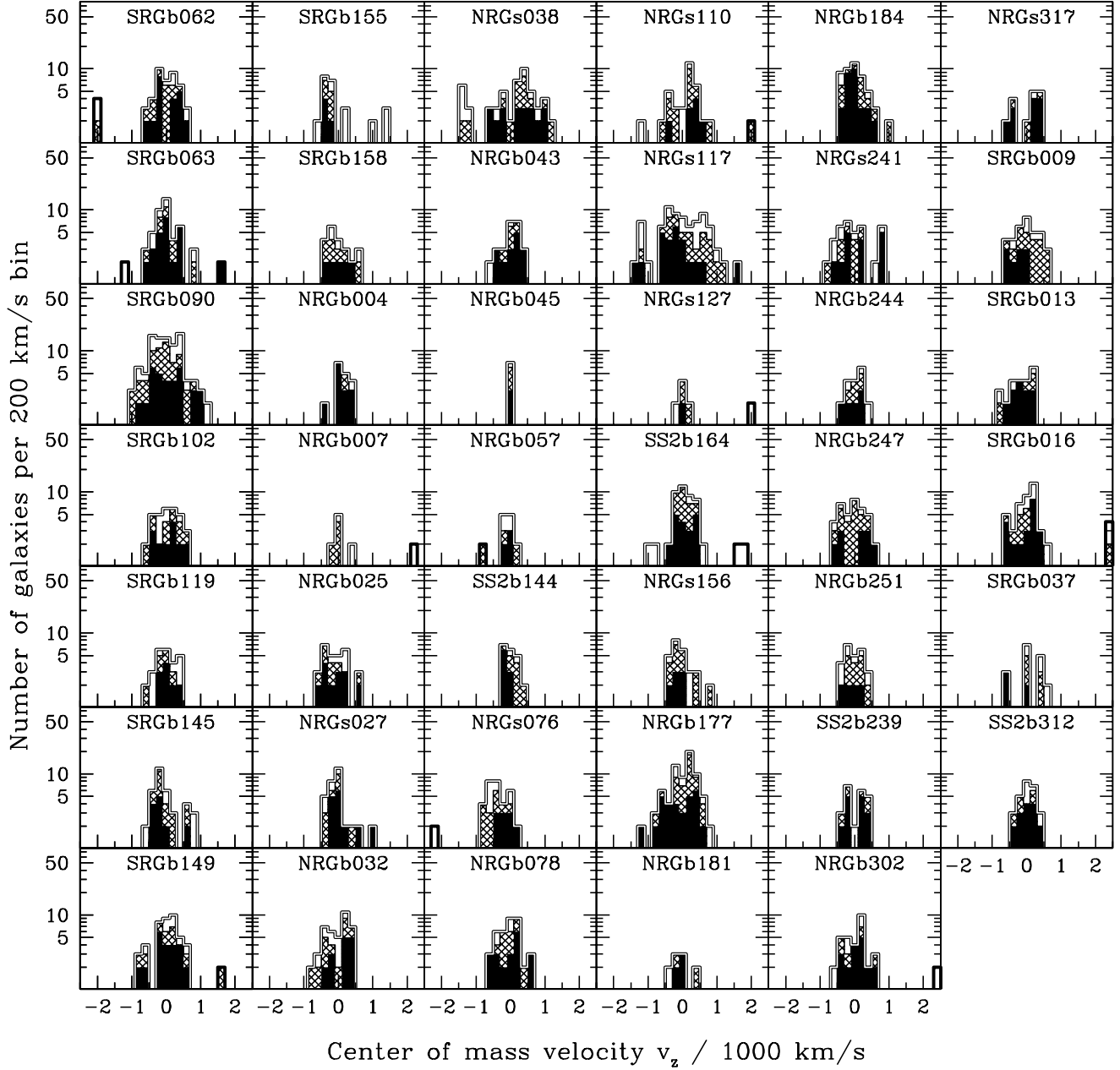


FIG. 1.— Velocity histograms of the galaxies observed in the 41 groups. Heavily outlined histograms indicate interlopers identified by the membership algorithm (§3.1). Filled, hatched, and empty histograms represent member galaxies within 0.5, 1.0, and 1.5 Mpc of the group center.

an extensive literature on the use of nonparametric adaptive kernels to reconstruct the cluster velocity distribution (Pisani 1993). However, the interpretation of the smooth probability distributions produced by the adaptive kernels is not straightforward. In particular, the frequent occurrence of slightly double-peaked probability distributions makes it difficult to tell exactly which galaxies should be considered cluster members. The Zabludoff et al. (1990) method is also nonparametric given a redshift range, and it reports the cluster membership unambiguously.

To implement this method, we begin with the known mean system velocity  $c\bar{z}$ , from Mahdavi et al. (2000) for groups, and from NED for the clusters. We remove all galaxies in our survey with  $|cz - c\bar{z}| > 3000 \text{ km s}^{-1}$  for each cluster. We then calculate  $\bar{z}$  and  $\sigma_z$ , given by

$$\sigma_z^2 = \frac{1}{N-1} \sum_{i=1}^N v_{z,i}^2, \quad (5)$$

where  $v_{z,i}$ , the center-of-mass velocity of each galaxy, is shown in equation (3). We sort the  $v_{z,i}$ , and remove all velocity-space neighbors with a velocity difference  $|\Delta v_z| > \sigma_z$ . We repeat the procedure until no galaxy is rejected. Tables 1 and 3 list the final membership counts. This list almost certainly contains additional interlopers, which we treat probabilistically in §4. A total of 1428 of the 2419 surveyed galaxies (60%) are group members. Of the 979 additional redshift taken from the literature, 851 (86%) are cluster members.

### 3.2. The Ensemble Systems

The data for any given system are scarce; there are on the average  $\sim 35$  members per system in the group database. It is therefore not possible to measure the shape of the matter density for individual systems. Our approach is to combine the groups and clusters into two ensembles. If the groups and clusters are self-similar—i.e., they are differently scaled versions of the same archetypal system—the aggregate yields the maximum information about the shape of the matter distribution. If on the other hand the systems in each ensemble are not self-similar, the ensemble itself represents an average over the various matter distributions, and the derived results ought to reflect the range of shapes present in each ensemble.

How are the systems to be combined? Ideally, both the velocity and the position of galaxies in each system ought to be scaled to a common value. We could choose a scaling such that the total virial mass of each system as given by equation (1) is the same:  $\tilde{v} \equiv v_z/\sigma_z$ , and  $\tilde{R} \equiv R/R_h$ . However, interlopers make the radial scaling problematic. In practice, scaling by  $R_h$  introduces noise in the data because of its sensitivity to interlopers and substructure (Mahdavi et al. 1999). A more suitable scaling radius is  $r_{200}$ , the radius at which the cluster density is 200 times the critical density of the universe.

A common method for calculating  $r_{200}$ , due to Carlberg et al. (1997), is to assume that the mass profile  $M(r) \propto r$  in the vicinity of  $r_{200}$ . Although observed X-ray and lensing cluster mass profiles are not proportional to  $r$  everywhere (Markevitch et al. 1999; Clowe et al. 2000), the linear relation does appear to be reasonable near  $r_{200}$ . This assumption, together with the virial theorem (with  $R_h$  replaced

by  $r_{200}$ ), yields the following adopted scalings:

$$\tilde{v}_z \equiv v_z/\sigma_z \quad (6)$$

$$r_{200} = \frac{\sqrt{3}\sigma_z}{10H_0(z)} \quad (7)$$

$$\tilde{R} \equiv R/r_{200} \quad (8)$$

where  $H_0(z)$  is the Hubble constant at the redshift of the cluster.

Using the velocity and radius scalings, we combine the groups and the clusters into two different ensembles—one “high  $\sigma$ ” and one “low  $\sigma$ ” ensemble. We do not combine all the data into one ensemble because lower mass systems of galaxies are thought to exhibit “similarity breaking,” or departures from the scaling laws that pertain to rich clusters of galaxies. For example Ponman et al. (1999), Helsdon & Ponman (2000), Mahdavi et al. (2000), and Mahdavi (2001) show that for low mass systems of galaxies, the relationships between the X-ray luminosity  $L_X$ , X-ray temperature  $T$ , and velocity dispersion  $\sigma$  differs from that of rich clusters. It would be interesting to see if corresponding differences in the matter distributions of these systems exist. Some high resolution N-body simulation have reported differences in the shape of the matter distribution for low- and high-mass systems (Jing & Suto 2000).

Another reason to break the data into two ensembles is completeness. In order for the analysis to be valid, the systems assigned to an ensemble need to have redshift surveys complete to the same scaled radius. Because of the fixed physical limits of our survey (1.5 Mpc), low- $\sigma$  (less massive) groups are complete to a much larger value of  $r_{200}$  than are high- $\sigma$  (more massive) systems. To make efficient use of the data, it is best to have two different ensembles with different completeness limits. For this purpose it is useful to define  $N(R_{\text{lim}})$  as the total number of galaxies within  $R_{\text{lim}}$  in all systems that are complete to  $R_{\text{lim}}$ . By maximizing  $N(R_{\text{lim}})$  as a function of  $R_{\text{lim}}$  for each of the two ensembles, we ensure that we are making the best possible use of the data while adhering to the completeness requirement.

Our goal is to have an equal number of galaxies in the low- and high- $\sigma$  ensembles. The closest we can come to this goal is to split the systems at  $\sigma = 470 \text{ km s}^{-1}$ . Maximizing  $N(R_{\text{lim}})$  yields a low- $\sigma$  ensemble with 893 members complete to  $R_{\text{lim}} = 1.8r_{200}$ , and a high- $\sigma$  ensemble with 945 members complete to  $R_{\text{lim}} = 0.8r_{200}$ . We therefore utilize 1838 (80%) of the 2279 available member galaxy velocities in our data sample. The two ensembles are shown in Figure 2.

### 3.3. Statistical Properties of the Ensembles

Table 5 shows the statistical properties of the final dynamical sample. We use the two sample KS test (Press et al. 1992) to evaluate the consistency of the high- and low- $\sigma$  ensembles. The KS tests show that the distributions of scaled velocities  $\tilde{v}_z$  in the two ensembles are consistent with each other. This is illustrated in the velocity histograms in Figure 3.

On the other hand, the distribution of scaled distances  $\tilde{R}_i$  differs in the two samples. According to the KS test, the probability that the distributions are the same is less than  $10^{-3}$ . Note that for this comparison only, we truncated the low- $\sigma$  sample at  $\tilde{R} = 0.8$ , because that is the

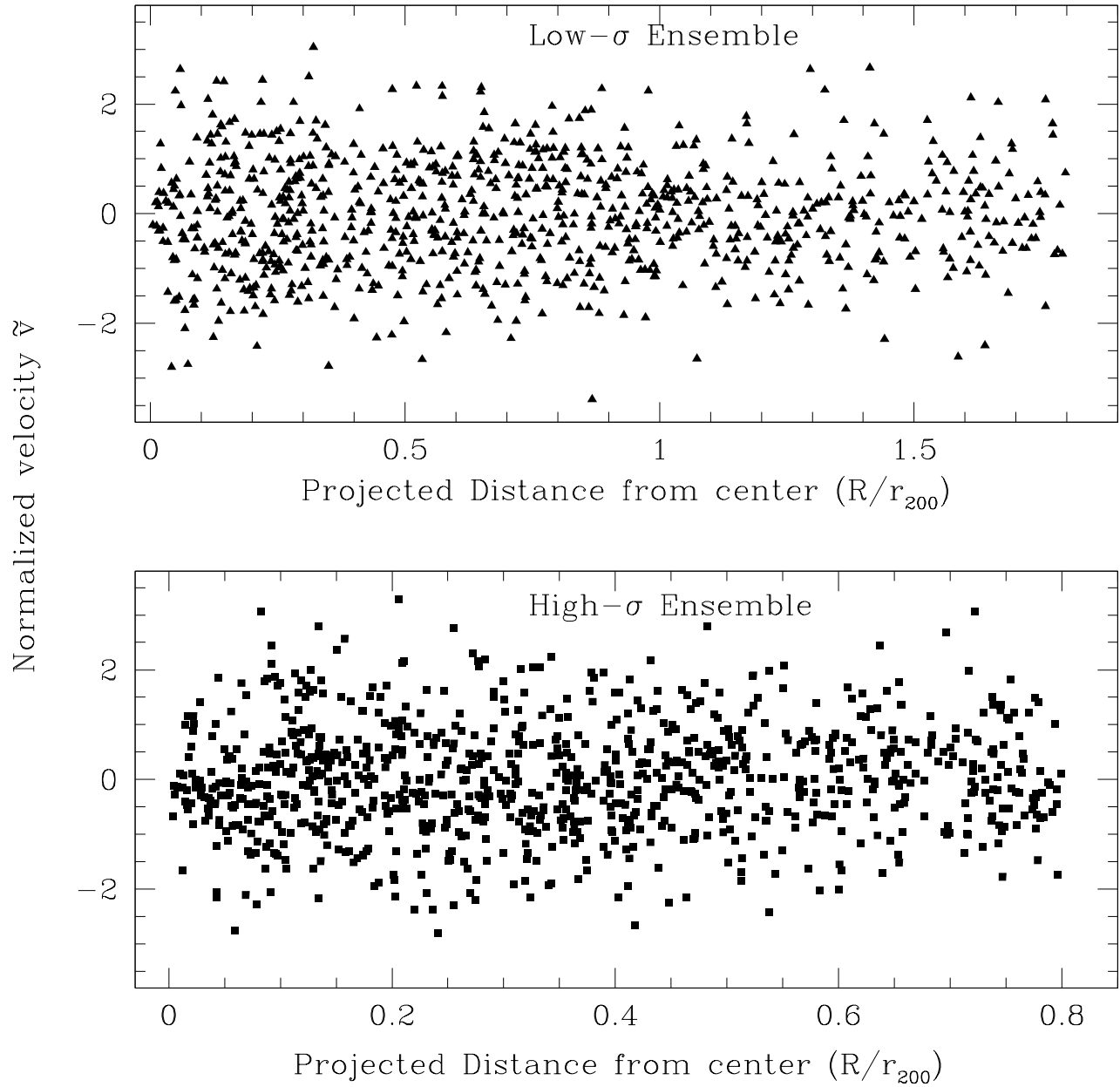


FIG. 2.— Phase space diagrams for the two ensembles.

limit to which the high- $\sigma$  sample is complete. Figure 3 also shows the  $\tilde{R}$  histogram. The distribution of the distances of high- $\sigma$  galaxies varies more rapidly with  $\tilde{R}$  than that of the low- $\sigma$  sample.

Thus, although the galaxies in the high- and low- $\sigma$  ensembles appear to have similar velocity distributions, their spatial distributions differ slightly. In the next section we investigate whether this difference translates into a difference in the total matter distribution.

#### 4. DYNAMICS

Here we discuss our technique for constraining the gravitational potential and orbital distribution of a system of galaxies. We use elements of Cuddeford (1991), Merritt & Saha (1993) and van der Marel et al. (2000) in our analysis. We add our own procedures for the faster computation of the phase space integrals described below.

##### 4.1. Measurement Technique

Our goal is to use measurements of galaxy positions and radial velocities ( $\mathbf{R}, \mathbf{v}_z$ ) to constrain the shape of the matter density. Our approach is Bayesian. For practical purposes, “Bayesian” means that our final result is a probability density rather a set of numbers and associated errors. For example, in addition to saying “the mass is  $M \pm \Delta M$  and the slope is  $n \pm \Delta n$ ,” we draw curves of equal probability in  $M - n$  space. These “confidence regions” will not necessarily be elliptical, as would be the case for a straightforward linearized  $\chi^2$  fit.

Bayes’s theorem specifies the method of calculating these confidence regions. Let the vector  $\mathbf{a}$  represent all the parameters we wish to derive from the data. We seek the probability of  $\mathbf{a}$  given the data, which, according to Bayes’s theorem, is<sup>4</sup>

$$p(\mathbf{a}|\mathbf{R}, \mathbf{v}_z) = \frac{p(\mathbf{a})}{p(\mathbf{R}, \mathbf{v}_z)} \prod_{i=1}^N p(R_i, v_{z,i}|\mathbf{a}) \quad (9)$$

The function on the left is the joint probability distribution of the unknown parameters  $\mathbf{a}$ , such that  $\int p(\mathbf{a}|R, v_z) d\mathbf{a} = 1$ . On the right hand side:

1.  $p(\mathbf{R}, \mathbf{v}_z)$  is independent of  $\mathbf{a}$ . Because the left hand side has to be normalized to unity, this probability is a precisely calculable number.
2.  $p(\mathbf{a})$  is the so-called Bayesian prior; it incorporates our prior expectations regarding the value of  $\mathbf{a}$ . The priors are detailed in §4.5 below.
3.  $p(R_i, v_{z,i}|\mathbf{a}) dR dv_z$  is the probability, given a physical model  $\mathbf{a}$ , of observing a galaxy with a projected distance from the center between  $R_i$  and  $R_i + dR$ , and with a line-of-sight velocity between  $v_{z,i}$  and  $v_{z,i} + dv_z$ .

Through Bayes’s theorem the desired confidence regions can be derived by calculating the two-dimensional function  $p(R, v_z|\mathbf{a})$  for all  $\mathbf{a}$  of interest, multiplying the result by the prior, and normalizing the answer to unity.

To do this calculation we use the galaxy phase-space distribution function (DF hereafter). The DF, a general

way to describe the dynamical state of the galaxies in a cluster, is sometimes written  $f(\mathbf{r}, \mathbf{v})$ , where  $\mathbf{r}$  and  $\mathbf{v}$  are the position and velocity vectors, such that  $f d^3\mathbf{r} d^3\mathbf{v}$  is the number of galaxies within a small phase space volume element  $d^3\mathbf{r} d^3\mathbf{v}$  at  $(\mathbf{r}, \mathbf{v})$ . If we knew the DF as a function of our parameter set  $\mathbf{a}$ , it would be simple to write down the probability of observing a galaxy with line-of-sight velocity  $v_z$  at a projected distance  $R$  from the cluster center. We begin by referring to this probability as  $p'(R, v_z|\mathbf{a})$ .

Suppose we conduct a complete redshift survey of an isolated spherical cluster complete to a limiting projected distance  $R_{\text{lim}}$  from the center. Then the probability of observing a given galaxy with radius  $R$  and velocity  $v_z$  is

$$p'(R, v_z|\mathbf{a}) = \frac{2\pi R}{N_0} \times \int \int \int f[\mathbf{a}(R, z, v_R, v_\phi, v_z)] dv_R dv_\phi dz \quad (10)$$

where the  $R$  and  $\phi$  denote polar coordinates in the plane of the sky, and  $z$  is the line of sight. The normalization  $N_0$  is the total number of galaxies within the limits of the survey, such that

$$\int_0^{R_{\text{lim}}} \int_{-\infty}^{\infty} p'(R, v_z|\mathbf{a}) dR dv_z = 1. \quad (11)$$

The chief difficulty lies in calculating the DF as a function of the dark matter potential and the galaxy orbital distribution. In other words, we must find the functions  $\mathbf{a}(R, z, v_R, v_\phi, v_z)$ . The probability  $p'$  represents an idealized observation of a cluster in the absence of contaminants (“interlopers”); the true probability  $p$ , which takes interlopers into account, is calculated below (§4.4).

We can begin by embedding this DF into our spherical gravitational potential  $\Phi(r)$ . Because  $\Phi$  is a negative quantity by convention, it is helpful to redefine  $\Psi \equiv \Phi(\infty) - \Phi(r)$ . Neglecting encounters, we assume that the only force on a galaxy is the Newtonian gravitational interaction  $\dot{\mathbf{v}} = \nabla\Psi$ . If we could write the DF in terms of the acceleration  $\dot{\mathbf{v}}$  and the potential  $\Psi$ , we would have the required link between the phase space variables and the gravitational potential. Such an equation would connect the observations to the matter distribution.

The Jeans theorem (Binney & Tremaine 1987) provides this crucial link. The theorem guarantees that in a steady-state spherical system the DF is a function of only two integrals of motion: the negative energy  $\mathcal{E} \equiv \Psi - v^2/2$ , where  $v^2 = \mathbf{v} \cdot \mathbf{v}$ , and the square of the angular momentum  $L^2$ . Thus we can write

$$f = f(\mathcal{E}, L^2) = f[\Psi(r) - v^2/2, |\mathbf{r} \times \mathbf{v}|^2] \quad (12)$$

Newton’s law and the Jeans theorem provide all the information we need to calculate the probability of observing a data point  $(R, v_z)$  given a DF and potential  $(f, \Psi)$  through equations (10)-(12). The Appendix contains the full expression for  $p'(R, v_z|\mathbf{a})$  in spherical coordinates.

Once we calculate the probability of the data given a specific model, we can use Bayes’s theorem to calculate constraints on the set of model parameters.

<sup>4</sup> For simplicity, we write the equations in this section in terms of  $R$  and  $v_z$ , rather than  $\tilde{R}$  and  $\tilde{v}_z$ .

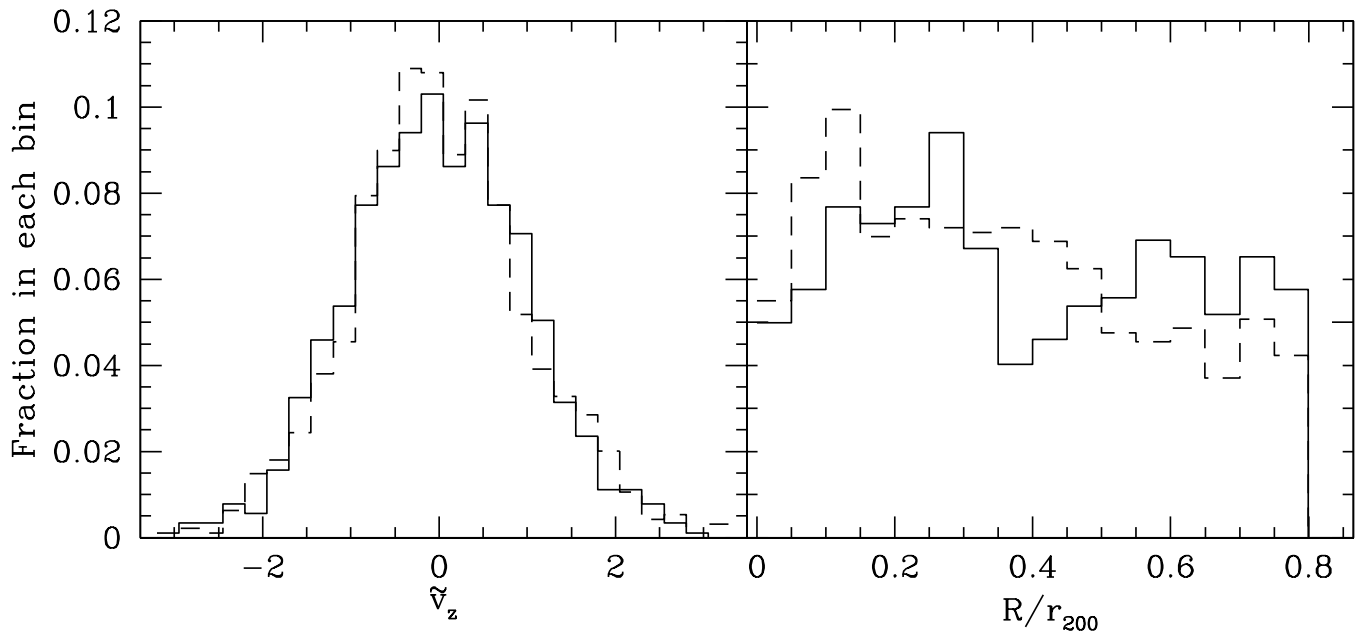


FIG. 3.— Comparison of the low- $\sigma$  sample (solid line) with the high- $\sigma$  sample (dashed line). Shown are the grand total velocity histogram (*left*) and the distribution of distances from the cluster center (*right*). For the latter plot, the low- $\sigma$  systems have been clipped at  $0.8r_{200}$ .

#### 4.2. Mass Model

To begin we consider the parameters of the mass profile. Many forms exist in the literature, but the ideal profile for our purposes should fulfill the following criteria, in decreasing order of importance: (1) It should be general enough to have the inner matter density slope as a free parameter; (2) it should include well-studied profiles as special cases; and (3) it should generate a density and gravitational potential in terms of elementary functions. This last property would make phase space distribution analysis faster and more straightforward.

We use a mass profile that fulfills all the three criteria:

$$M(r) = M_0 \left( \frac{r}{r+r_0} \right)^{3-n}, \quad (13)$$

where  $M_0$  is the total mass of the system,  $r_0$  is the characteristic radius, and  $n$  is the inner slope of the matter density. This model is sometimes referred to as the “gamma-model” (not to be confused with  $\gamma$  below) in stellar-dynamical literature (Dehnen 1993; Tremaine et al. 1994). The density and potential pair generated by this mass profile are

$$\rho(r) = \frac{M_0 r_0 (3-n)}{4\pi} r^{-n} (r_0 + r)^{n-4} \quad (14)$$

$$\Phi(r) = -\frac{GM_0}{r_0(2-n)} \left[ 1 - \left( \frac{r}{r+r_0} \right)^{2-n} \right] \quad (15)$$

The above profile includes the well-known Hernquist (1990) and Jaffe (1983) models, which are generated by the  $n = 1$  and  $n = 2$  cases, respectively. The Hernquist profile is known to match the light and matter distribution in elliptical galaxies as well as in groups and clusters (Mahdavi et al. 1999; Rines et al. 2001). The formulation above is a generalization with arbitrary inner slope  $n$ . Note that the NFW model  $\rho_{\text{NFW}} \propto r^{-1}(1+r)^{-2}$  has infinite mass

and is not a subset of the models we consider. Using a method completely different from ours, Rines et al. (2003) find that samples as large  $10^4$  galaxies cannot distinguish between the NFW and Hernquist mass distributions in the infall regions of clusters ( $R > 2$  Mpc).

As a further boon, the generalized potential is invertible analytically:

$$r(\Phi) = \frac{r_0}{1 - [1 + (n-2)\Phi r_0/GM_0]^{1/(2-n)}}, \quad (16)$$

further simplifying the calculation of phase space densities described below.

In summary, we begin with a three dimensional parameter set  $\mathbf{a}$ , containing the inner slope, the transition radius, and the normalization of the matter density. Three more parameters not directly related to the matter density are also required. They are the slope of the galaxy density at infinity,  $\gamma$ , the velocity anisotropy,  $\beta$ , and the interloper fraction,  $P_I$ .

#### 4.3. DF model

Next we examine the form for the DF,  $f(\mathcal{E}, L^2)$ . It is virtually a guarantee that given a potential  $\Psi$ , many quite different  $f$  will describe the same data equally well. We therefore use a very simple DF—a scale free, separable function of the energy and angular momentum (Fricke 1952):

$$f(\mathcal{E}, L^2) = f_0 \mathcal{E}^{\alpha-1/2} L^{-2\beta}. \quad (17)$$

As we show below, this well-studied DF adequately describes the cluster and group data.

This DF yields a unique galaxy density profile  $\nu$  (distinct from the matter density profile  $\rho$ ). To calculate a galaxy density profile, we note that the integral of this particular DF over all velocities gives the particle space



density; see the Appendix for a detailed calculation of the integral:

$$\nu(r) = \int_{v^2 \leq 2\psi} f(\mathcal{E}, L^2) d^3\mathbf{v} \quad (18)$$

$$\propto r^{-2\beta} \Psi(r)^{\alpha-\beta+1}. \quad (19)$$

This  $\beta$  parameter is equal to the orbital anisotropy of the galaxies:

$$\beta = 1 - \frac{\sigma_t^2}{\sigma_r^2} \quad (20)$$

In other words,  $\sqrt{1-\beta}$  is the ratio of the velocity dispersion in the tangential direction to the velocity dispersion in the radial direction. From both the equations above it is evident that  $\beta < 1$ . If that were not the case, (1) the galaxy mass  $\int r^2 \nu(r) dr$  would diverge at  $r = 0$ , and (2) one of the velocity dispersions  $\sigma_t$  and  $\sigma_r$  would have to be imaginary.

The  $\alpha$  and  $\beta$  parameters also describe the slope of the galaxy density at infinity. For example, if  $\Psi(r) \propto 1/r$  at large radii, then  $\nu(r) \propto r^{-\gamma}$ , with

$$\gamma = \alpha + \beta + 1. \quad (21)$$

Henceforth, without loss of generality, we will discuss the energy slope of the DF solely in terms of  $\gamma$ . As  $\gamma$  increases, the cluster contains fewer galaxies with large kinetic energies. This behavior occurs because galaxies with energy  $\mathcal{E} \sim \Phi$  begin to outnumber galaxies with  $\mathcal{E} \approx 0$  (see equation 12). As a result the galaxy distribution is less extended than it would be with small  $\gamma$ . The minimum acceptable value is  $\gamma = 4$ , because the mass models we consider (equation 14) decline asymptotically as  $r^{-4}$ .

The inner slope of the galaxy density  $n_g$  can be different from the slope of the matter density  $n$ :

$$n_g = \begin{cases} 2\beta & (n < 2) \\ 2\beta + (n-2)(\gamma-2\beta) & (n > 2) \end{cases} \quad (22)$$

The reason for the split at  $n = 2$  is that the potential  $\Phi$  becomes singular for  $n \geq 2$ , and through equation (19), gives a contribution to the inner slope of the galaxy density. For  $n < 2$ , the potential is constant at  $r = 0$ , and there is no such contribution. A corollary of this property is the requirement

$$n > n_g. \quad (23)$$

This statement specifies that the galaxy density may never exceed the total matter density. It implies, importantly, that for constant density cores  $n = 0$ , no radially anisotropic models ( $\beta > 0$ ) are viable. This requirement is not unique to our model; a large number of other radially anisotropic models, with substantially different DFs and anisotropy profiles  $\beta(r)$ , are also incompatible with constant-density cores (van der Marel et al. 2000; Cuddeford 1991).

Up to this point we have a five-dimensional parameter space  $\mathbf{a} = (M_0, n, r_0, \gamma, \beta)$ . One final dimension is necessary, the interloper fraction.

#### 4.4. Interlopers

So far we have assumed that the clusters and groups we observe exist in isolation, free from any contamination from background and foreground galaxies. In reality, however, even after velocity clipping (see §3.1), a small

percentage of the galaxies are unrelated objects projected into the phase space volume of interest. We wish to estimate this percentage—the interloper fraction  $P_I$ —from the data.

We follow the elegant method of van der Marel et al. (2000) in treating the interlopers statistically. This method assumes that interlopers land randomly in the group or cluster phase space volume. Thus every galaxy we observe has a finite probability  $P_I$  of being an unrelated interloper. The probability that a galaxy with velocity  $v_z$  and projected distance  $R$  is included in our survey is then

$$p(R_i, v_{z,i} | \mathbf{a}) = (1 - P_I) p'(R_i, v_{z,i} | \mathbf{a}) + \frac{2P_I R_i}{R_{\text{lim}}^2 |v_{\text{lim}}|} \quad (24)$$

Where  $|v_{\text{lim}}|$  is the maximum observed line of sight velocity. It is easy to verify that the above expression meets a normalization criterion similar to equation (11) as long  $p'(R, v_z | \mathbf{a}) \rightarrow 0$  as  $v_z \rightarrow |v_{\text{lim}}|$ :

$$\int_0^{R_{\text{lim}}} \int_{-|v_{\text{lim}}|}^{|v_{\text{lim}}|} p(R, v_z | \mathbf{a}) dR dv_z = 1. \quad (25)$$

We can now write down the final expression for the six-dimensional function  $p(\mathbf{a} | \mathbf{R}, v_z)$ . The full expression is too long to be listed here and is derived in detail within the Appendix. The derivation consists of a straightforward if laborious combination of equations (10) - (24). The main complications occur because we observe a spherical object in cylindrical coordinates.

We henceforth refer to the joint probability distribution of observables for a given measurement simply as  $p(\tilde{R}, \tilde{v}_z)$ , where without loss of generality we have made the substitution  $R \rightarrow R/r_{200} \equiv \tilde{R}$  and  $\tilde{v}_z \rightarrow v_z/\sigma_z \equiv \tilde{v}_z$ .

#### 4.5. Reparameterization and Bayesian Priors

The six parameters of our model are  $(n, M_0, \gamma, r_0, \beta, P_I)$ . Some of these parameters can be redefined without loss of generality to yield more compact joint confidence regions. For example, we find that jointly constraining  $n$ ,  $M_0$  and  $r_0$  as defined in equation (13) yields a high degree of correlation among all three parameters. However, by switching to the system

$$r_{5/2} \equiv \frac{3r_0}{5-2n}, \quad (26)$$

$$M_{200} \equiv M(r_{200}), \quad (27)$$

we significantly reduce the overall extent of the confidence regions. The new radius scaling  $r_{5/2}$  is defined such that

$$\left. \frac{d \ln \rho}{d \ln r} \right|_{r=r_{5/2}} = -\frac{5}{2} \quad (28)$$

In other words,  $r_{5/2}$  is the location at which the slope of the matter density is 5/2. Similarly,  $M_{200}$  is the total mass contained within  $r_{200}$ .

The fact that we analyze ensembles of clusters has implications for the measurement of  $r_{5/2}$  and  $M_{200}$ . To create the ensembles, it was necessary to scale the radii and line-of-sight velocities by  $r_{200}$  and  $\sigma$ , respectively (see §3.2). Because of this scaling, the mass and characteristic radius of the ensembles are normalized as well. As a result, our method constrains the dimensionless quantity

$$\tilde{M}_{200} \equiv \frac{GM_{200}}{r_0 \sigma^2} \quad (29)$$

for each of the two the ensembles. Similarly, the quantity

$$\tilde{r}_{5/2} \equiv r_{5/2}/r_{200}, \quad (30)$$

rather than  $r_{5/2}$  alone, is constrained by our method.

A redefinition of the anisotropy is also appropriate. Because the allowable range in  $\beta$  is  $[-\infty, 1]$ , we define, similarly to Wilkinson et al. (2002),

$$\tilde{\beta} \equiv -\ln(1 - \beta). \quad (31)$$

This redefinition maps the allowable range from  $[-\infty, 1]$  into  $[-\infty, \infty]$ , and ensures that the parameter space has equal prior probability density in the radially and tangentially anisotropic regimes.

The final parameter set is  $(n, \tilde{M}_{200}, \tilde{r}_{5/2}, \gamma, \tilde{\beta}, P_I)$ . In our analysis, we take the Bayesian priors  $p(\mathbf{a})$  (equation 9) for these parameters to be uniform.

#### 4.6. Computation Technique

Here we describe the numerical procedure used to calculate our probabilities. Even though  $p(\tilde{R}, \tilde{v}_z)$  has six dimensions, we limit ourselves to constraining five at any given time. As we explain in §5.2, for free  $n$  our data can place only a lower, and not an upper limit on the energy slope  $\gamma$ , and therefore  $p(\tilde{R}, \tilde{v}_z)$  cannot be normalized properly as a function of  $\gamma$ . On the other hand, if we fix  $n$ , then we can easily constrain  $\gamma$ . For this reason, we conduct the entire procedure described below several times: (a) once with  $n$  fixed and  $\gamma$  free (§5.1), and (b) many times with  $n$  free and  $\gamma$  ranging over several discrete values (§5.2).

We use a computation technique that yields an estimate of  $p(\tilde{R}, \tilde{v}_z)$  on an grid with  $m^5$  elements:

1. *Tabulation.* There are two scaling parameters, the transition radius  $r_{5/2}$  and the mass normalization  $M_{200}$ . We define two dimensionless parameters,  $R' \equiv \tilde{R}/\tilde{r}_{5/2}$ , and  $v'^2 \equiv \tilde{v}_z^2 r_0 / (GM_{200})$ . Then we tabulate  $p(R', v' | \mathbf{a})$  on a refined  $M \times M$  grid for different values of  $\tilde{\beta}$ , and  $n$ . This step therefore requires  $m^2 M^2$  evaluations of the triple integral given in equation (10). This phase occupies roughly 25% of total computation time.
2. *Interpolation.* For each of the  $N_0$  data points  $(\tilde{R}_i, \tilde{v}_{z,i})$ , we use cubic spline interpolation to calculate the probability density as a function of  $\tilde{\beta}$ ,  $\tilde{M}_{200}$ ,  $\tilde{r}_{5/2}$ , and either  $n$  or  $\gamma$ . This step requires  $N_0 m^4$  interpolations, and takes up roughly 70% of the total computation time. Because of the linearity of the interloper parameter in equation (24), calculating the probability density as a function of  $P_I$  takes up a trivial amount of time in comparison to the  $N_0 m^4$  interpolations.
3. *Projection.* To display the 5-dimensional probability density, we produce two-dimensional realizations of it. This means creating  $5 \times 4/2 = 10$  separate probability distributions, each representing a summing (or “marginalization”) of  $p(\mathbf{a} | \mathbf{R}, \mathbf{v}_z)$  along a cube orthogonal to the other 9. For example, to obtain the joint probability distribution of the mass normalization  $\tilde{M}_{200}$  and inner slope  $n$ , we use

$$\int \int \int p(\tilde{M}_{200}, n, \tilde{r}_{5/2}, \tilde{\beta}, P_I) d\tilde{r}_{5/2} d\tilde{\beta} dP_I. \quad (32)$$

This step is conducted simultaneously with step (2) and takes up a negligible amount of the total computation time.

4. *Goodness of fit.* To check the goodness-of-fit of the physical model, we bin the data in two dimensions and use  $\chi^2$  tests to compare the number of galaxies in each bin to the number predicted by the most likely model. We bin the data only to check the goodness of fit, not to derive constraints on the parameters. We calculate the quantity

$$X^2 \equiv \sum_{j=1}^k \frac{(N_j - p_j N)^2}{N_j}, \quad (33)$$

where  $k$  is the total number of bins,  $N_j$  is the number of galaxies observed in bin  $j$ , and  $p_j$  is the integrated probability of the best fit model in that bin. The quantity in the denominator is the Poisson variance in each bin.

It can be shown that for our case  $X^2$  is bounded above by a  $\chi_{k-1}^2$  distribution and below by a  $\chi_{k-s-1}^2$  distribution (Lupton 1993), where  $s = 5$  is the number of parameters we fit. Therefore, if  $1 - q$  is confidence with which we can reject the hypothesis that the data is consistent without our model, then  $q$  must lie between  $\Gamma[(k - s - 1)/2, X^2/2]$  and  $\Gamma[(k - 1)/2, X^2/2]$ , where  $\Gamma[a, x]$  is the incomplete gamma function. We always use the value that yields the smaller  $q$ .

We create rectangular bins in  $R, v_z$  space so that they each contain 30 galaxies and so that  $\sum p_j = 1$ . Let  $t$  be the largest integer smaller than or equal to  $\sqrt{N_0/30}$ . There are  $t$  strips in the radial direction, each with 30 galaxies. The first  $t - 1$  radial strips are divided into  $t$  bins each containing 30 galaxies; the last radial strip is divided into  $t - 1$  bins each containing 30 galaxies, and a final bin containing the remaining galaxies. Therefore, the total number of bins is  $k = t^2$ .

For our final calculations, we use  $M = 50$  and  $m = 40$ .

#### 4.7. Simulations

The ideal procedure to test the validity of our technique would be to create thousands of Monte Carlo realizations of various clusters with different values of  $\beta$ ,  $\gamma$ ,  $r_0$ ,  $M_{200}$ , and  $P_I$ , and to conduct our analysis separately for each realization. Then we could compute the fraction of simulations in which the most likely parameters closely match the input parameters. Unfortunately, such a test would be prohibitively expensive, requiring many months of CPU time.

As a compromise, we conduct four tests in which we attempt to model simulated ensembles. In these simulations, we draw 893 members from the  $p(\tilde{R}, \tilde{v}_z)$  for a set of parameters given *a priori*. The procedure for drawing deviates distributed according to  $p(\tilde{R}, \tilde{v}_z)$  is the rejection method described in Press et al. (1992). The rejection method uses a comparison function that is everywhere

greater than  $p(\tilde{R}, \tilde{v}_z)$  and has an analytic cumulative distribution. A large number of deviates are drawn from the comparison function, and those that lie beneath  $p(\tilde{R}, \tilde{v}_z)$  have the desired probability distribution. We extend the Press et al. (1992) technique to two dimensions by using a constant comparison function in  $R, \tilde{v}_z$ . Finally, we maximize the Bayesian likelihood by conducting the entire procedure described in the previous section. The simulated data are drawn from models with  $\gamma = 8$ . We first analyze the simulations assuming  $\gamma = 8$ , and then we consider the effects of varying  $\gamma$ .

Table 4 lists the parameters used for the simulations, and Figure 4 shows their properties. The phase space distribution generated from each parameter set appears together with the grand total velocity histogram,

$$N(\tilde{v}_z) \equiv \int_0^{\tilde{R}_{\text{lim}}} p(\tilde{R}, \tilde{v}_z) d\tilde{R} \quad (34)$$

and the surface number density

$$\Sigma(\tilde{R}) \equiv N_0 \int_{-\infty}^{\infty} p(\tilde{R}, \tilde{v}_z) d\tilde{v}_z / (2\pi\tilde{R}), \quad (35)$$

where  $N_0$  is the total number of galaxies in the sample. In these figures we also show the most likely models for the above two observables. Note that these curves are not fits, but predictions of the unbinned maximum likelihood technique described in the previous section.

According to the simulations, our estimation technique is able to tell apart fundamentally different models with similar binned statistics. For example, simulation A and D have similar  $N(\tilde{v}_z)$  and  $\Sigma(\tilde{R})$ , even though A was generated with  $n = 1$  and D was generated with  $n = 0$ . Because we analyze unbinned velocities and radii, our technique yields a 95% confidence region  $n = 0.6\text{--}1.3$  for A and  $n = 0\text{--}0.7$  for D. The unbinned estimator is able to distinguish the parent distributions of the two data sets.

Another feature of the simulations is that they are more powerful when  $n$  is steep—that is, the slope of the matter density is large. In this case our method yields a fairly accurate measurement of  $n$ . When  $n$  is close to 0—that is, the matter density has a flat core—then the measurement method yields a wider range in  $n$ . Nevertheless, this range for both models C and D includes the correct  $n = 0$  solution.

We can also use the simulations to test the validity of our numerical integration. The integrals of  $N(\tilde{v}_z)$  and  $2\pi\tilde{R}\Sigma(\tilde{R})/N_0$  should equal 1. For the four simulations, we find that  $\int N(\tilde{v}_z) = 0.9976, 0.9880, 0.9998, 0.9987$ , and  $\int 2\pi\tilde{R}\Sigma(\tilde{R})/N_0 = 0.9987, 0.9972, 1.0001, 0.9995$ . Thus in most cases our accuracy is better than 0.5%.

Finally, it is useful to examine how sensitive the measurement technique is to variations in  $\gamma$ . In analyzing our simulations, we have set  $\gamma = 8$ , i.e., we have fixed it at the value used to generate the data. Now we reanalyze the data assuming different values of  $\gamma$ , from 4 to 12. We rerun our analysis procedure, calculating the quality of fit  $q$  and the 95% marginalized confidence region in  $n$  as a function of  $\gamma$ . The results are shown in Figure 5. We find that as long as  $q > 0.1$ , i.e., the fit is statistically acceptable, then the derived confidence regions  $n_{\text{fit}}$  are close to the true  $n$  for most values of  $\gamma$ . In general, choosing the *smallest value* of  $\gamma$  that gives  $q > 0.1$  results in the most

accurate confidence intervals on  $n$ . Choosing larger values of  $\gamma$  yields confidence regions  $n_{\text{fit}}$  that are slightly different from the correct values.

While not exhaustive, these simulations suggest that our measurement technique is trustworthy. The 95% confidence contours derived using our method should include the correct value of each parameter if the general model is admissible (i.e.,  $q \gtrsim 0.1$ ), with the caveat that the smallest value of  $\gamma$  that yields a good fit should be used. These confidence regions ought to be compact for large values of the density slope  $n$ , and broad-valued for flat halos with small  $n$ .

## 5. RESULTS

Here we describe our constraints on the inner matter density slope  $n$ , the mass normalization  $\tilde{M}_{200}$ , the transition radius  $\tilde{r}_{5/2}$ , the velocity anisotropy  $\tilde{\beta}$ , and the interloper fraction  $P_I$ . It bears repeating that our constraints result from maximization of the likelihood function described above, using the full unbinned data set. The binned profiles we show are for illustrative purposes and do not indicate ordinary  $\chi^2$  fitting.

### 5.1. Models with Constant-Density Cores

First we investigate the possibility that the ensembles have constant-density cores ( $n = 0$ ). In §4.3 we showed that such cores do not support radially anisotropic models; we are limited to  $\beta \leq 0$ . Therefore the minimization occurs over five dimensions:  $(\tilde{\beta}, \gamma, \tilde{r}_{5/2}, \tilde{M}_{200}, P_I)$ .

The dotted line in figure 6 shows the results of the minimization. We show the most likely grand total velocity histogram and galaxy surface density as given by equations (34)–(35).

We find that matter profiles with constant-density cores do not produce a galaxy density profile as steep as our data. Although the grand-total velocity histogram  $N(\tilde{v}_z)$  seems accurately reproduced, the model's galaxy density in the innermost regions is too small. The constant density models are disfavored with  $q = 0.003$  for the low- $\sigma$  ensemble, and  $10^{-5}$  for the high- $\sigma$  ensemble.

Note that the inconsistency with the galaxy surface density  $\Sigma$  is not the only discrepancy between the data and the  $n = 0$  model. Our  $\chi^2$  goodness-of-fit method described above shows that the  $n = 0$  models predict too many high-velocity galaxies at intermediate radii. This discrepancy is not easily visualized through the plots of  $N(\tilde{v}_z)$  or  $\Sigma(\tilde{R})$ .

Just to be sure that the  $n = 0$  models are a poor fit, we also try relaxing the requirement that the orbits not be radially anisotropic. We fit the full range of negative to positive  $\beta$ , and find that  $\beta \approx 0.25, \gamma \approx 4$  maximize the likelihood for both ensembles. These values are formally unphysical, yielding a galaxy density that is steeper than the  $n = 0$  total matter density near the central regions. However, even these fits are rejected with  $q = 0.02$  for groups and  $q = 0.003$  for clusters.

It is interesting to note (Figure 6) that the inconsistency occurs near  $r = 0$ , and not at larger radii. If our choice of mass model (equation 14) were the cause for the poor fit, we would expect the inconsistency to occur at the larger radii, because that is where we impose a  $\rho \propto r^{-4}$  behavior. We conclude that models with a constant density core ( $n = 0$ ) are at best barely consistent with the data.

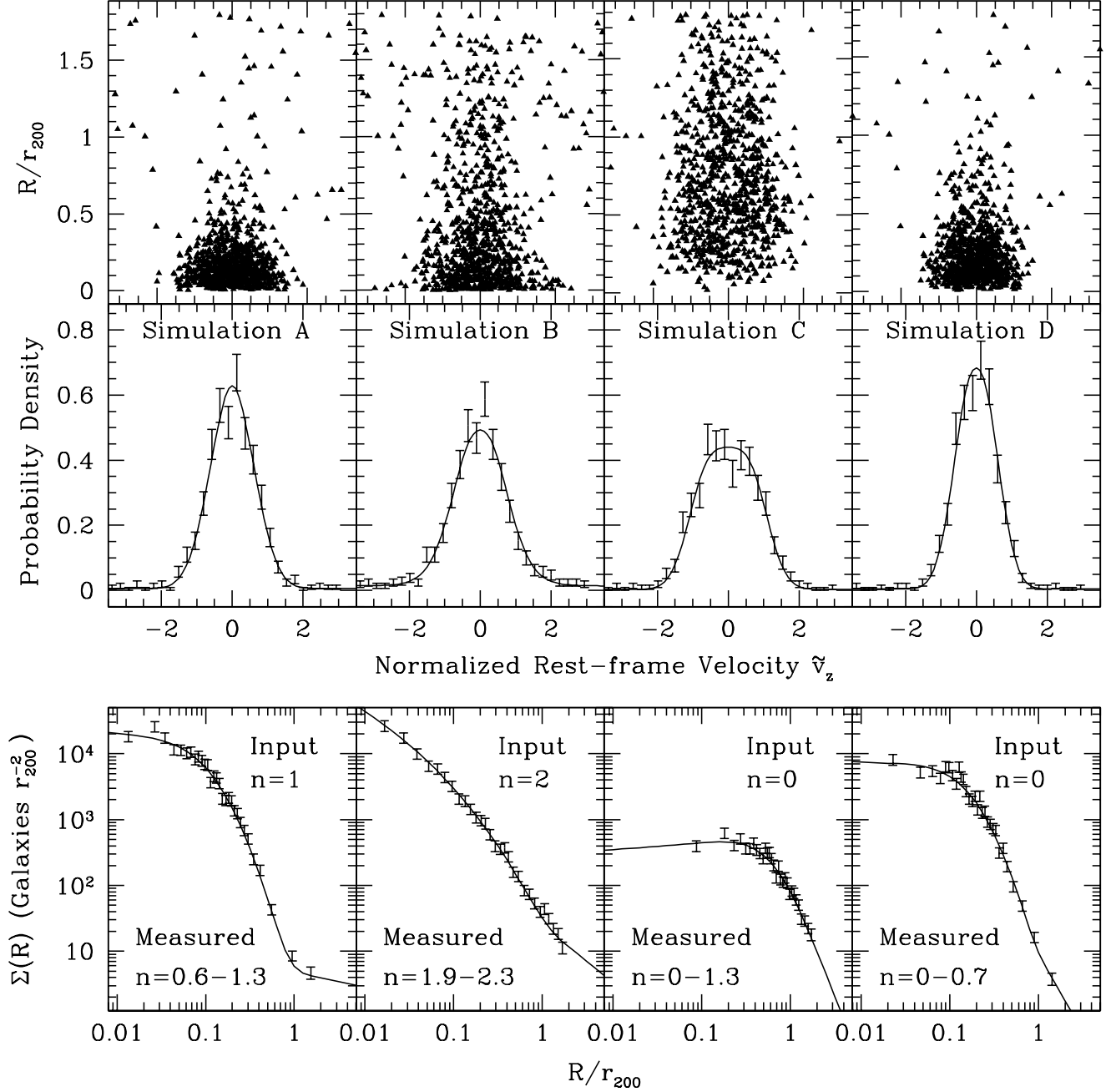


FIG. 4.— Simulations of the technique: (*top*) phase space distributions generated using the four parameter sets in Table 4; (*middle*) the grand total velocity histogram  $N(\tilde{v}_z)$ ; (*bottom*) the galaxy surface density. The solid lines show predictions (not fits) of the maximum-likelihood analysis of unbinned distance and velocity data. The “breaks” in  $\Sigma(R)$  around  $R \sim r_{200}$  in simulations A and B are due to contamination by the interloper population  $P_I$ .

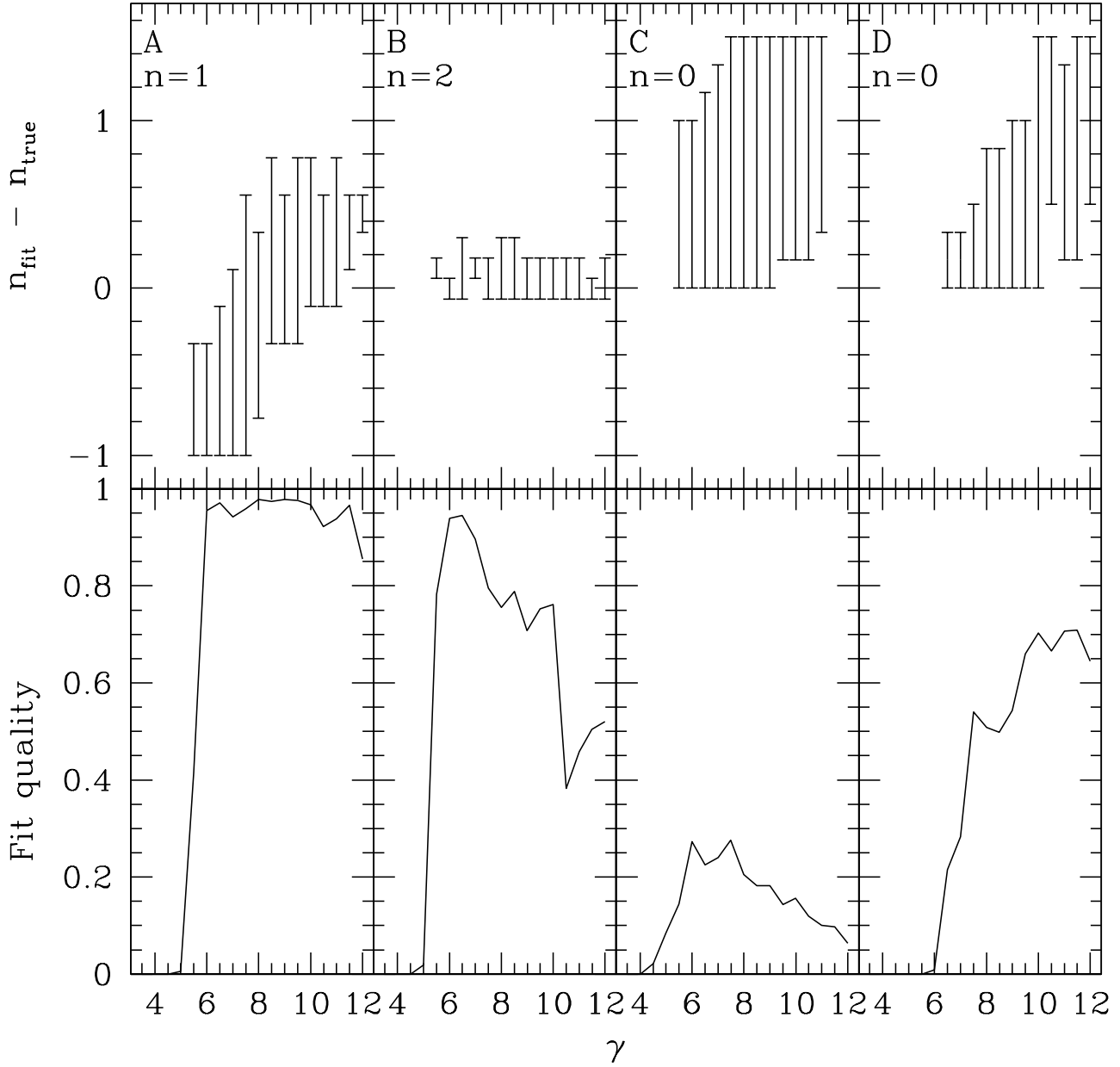


FIG. 5.— Testing the robustness of the simulations as a function of  $\gamma$ : (*top*) the 95% confidence regions on the best-fit  $n$  minus the true  $n$  for each simulation; (*bottom*) the quality of fit as a function of  $\gamma$ .

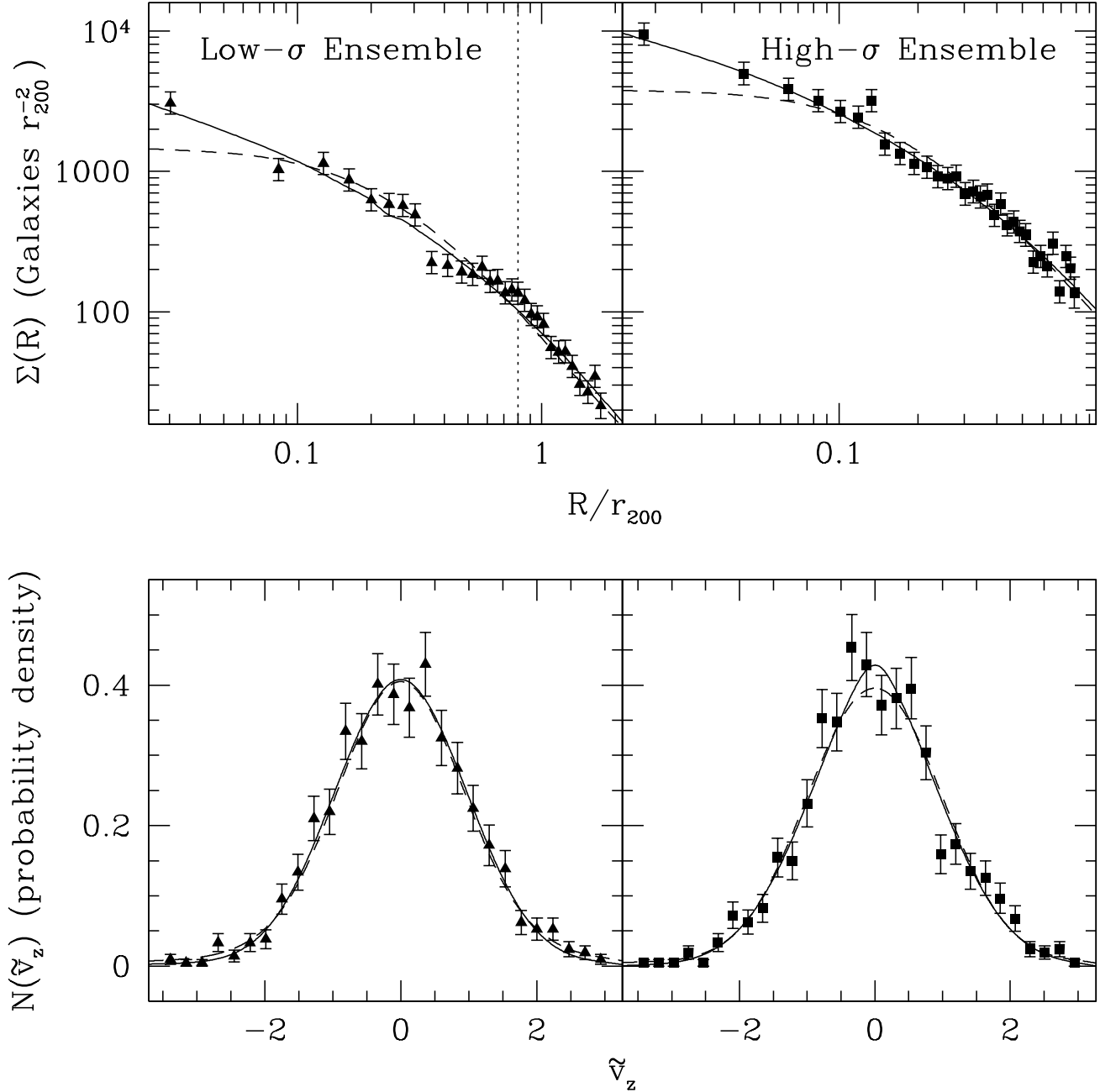


FIG. 6.— Predictions (not fits) of the maximum-likelihood analysis of unbinned distance and velocity data. Shown are the surface number density (*top*) and the grand total velocity histogram (*bottom*). The solid line represents the most likely model overall, whereas the dashed line shows the most likely model when the inner slope of the matter density is forced to equal 0. The latter model is rejected with 99.7% or better confidence. The dotted vertical line represent the completeness limit of the high- $\sigma$  ensemble.

### 5.2. Models with Fixed $\gamma$

Now we allow the inner slope of the mass density  $n$  to vary, maximizing  $p(\tilde{R}, \tilde{v}_z | \mathbf{a})$  for several different values of  $\gamma$ , the slope of the galaxy density at large radii. The jointly constrained parameter set is therefore  $(n, \tilde{\beta}, \tilde{r}_{5/2}, \tilde{M}_{200}, P_I)$ . As described in §4.3,  $\gamma$  is also related to the slope of the energy term in the distribution function (equation 12). The larger the value of  $\gamma$ , the more peaked the distribution of galaxies with small kinetic energies.

Our data can provide a lower, but not an upper limit on  $\gamma$ . For both the low- and the high- $\sigma$  ensembles, the quality of the fit increases with  $\gamma$ , asymptotically approaching a fixed value as  $\gamma \rightarrow \infty$ . Our simulations (§4.7) show a similar effect, and suggest that the smallest value of  $\gamma$  that gives an acceptable fit is likely to give the most accurate results. Therefore, instead of constraining  $\gamma$  continuously, we list the minimum value of  $\gamma$  that yields a quality of fit  $q > 0.1$ . This constraint corresponds to  $\gamma = 8$  for the low- $\sigma$  and  $\gamma = 12$  for the high- $\sigma$  ensemble.

Figures 6 and 7 show the results. Aside from difference in the value of  $\gamma$ , the high- and low- velocity dispersion systems provide very similar fits: the galaxy orbits are consistent with isotropic ( $\tilde{\beta} = 0$ ), and both ensembles have  $n \approx 2$  and transition radius  $r_{5/2}$  much greater than  $r_{200}$ . In other words, the matter distributions for both samples are consistent with a singular isothermal sphere within the limits of the survey.

### 5.3. Discussion

Neither the cluster nor the group data supports a total matter distribution with a constant-density core ( $n = 0$ ). This result is in keeping with recent studies of more distant clusters in the X-ray (Lewis et al. 2003; Arabadjis et al. 2002; Allen et al. 2002) as well as weak lensing maps (Gavazzi et al. 2003; Clowe & Schneider 2002; Clowe et al. 2000). Optical observations of cluster velocity data have also yielded similar results. For example, Biviano & Girardi (2003) uses the Jeans equation and isotropic velocity dispersion profiles to rule out constant-density cores in the Two Degree Field Galaxy Redshift Survey (Colless et al. 2001).

Our results suggest that within a projected region  $R \lesssim 2r_{200}$ , the matter distribution in both low- and high- $\sigma$  systems of galaxies is close to a single power law with  $n = 2$ . The density declines rapidly only outside  $2r_{200}$ . This result is consistent with N-body simulations. For example, in the original paper describing the simulations of NFW, the density profiles of dark matter halos are consistent with a single power law of slope 2 between  $0.1r_{200}$  and  $2r_{200}$  (see their Figure 3). The only discrepancy, or flattening, occurs at radii smaller than  $0.1r_{200}$ , where our datasets have 56 (low- $\sigma$ ) and 131 (high- $\sigma$ ) members, perhaps insufficient to provide a robust constraint.

A major limitation of our method is the fixed form of the galaxy distribution function  $f(\mathcal{E}, L^2)$ . By modeling the DF as a power law in energy and angular momentum, we neglect all other possible forms for the DF, some of which may indeed be consistent with an  $n = 0$  model. On the other hand, the  $\chi^2$  tests we perform reassure us that the data are at least statistically consistent with cuspy matter distributions.

The work of van der Marel et al. (2000), whose methods

we adapt and extend, involves a similar analysis, and it is instructive to compare the two different approaches. Instead of fixing the form of the DF, they fix the form of the surface number density  $\Sigma(R)$ , and calculate the DF using an inversion of the integral in equation (18). As a result, their DF is more general than ours,  $f \propto g(\mathcal{E})L^{-2\beta}$ , with  $g$  being the result of the inversion. Their method has the advantage of generality, but the disadvantage that it requires a two-stage fit: first  $\Sigma(R)$  must be fit and fixed, and then  $\beta$  and  $n$  measured separately, without an indication of how changes in the  $\Sigma(R)$  fit would affect the resulting constraints. Our method sacrifices generality by setting  $g(\mathcal{E}) = \mathcal{E}^{\alpha-1/2}$ , but gains the advantage that all but one of the parameters are fit simultaneously ( $\gamma$  is varied discretely). As a result, the correlation among the parameters is easy to understand via marginalized confidence contours (Figure 7).

Another difference between our work and van der Marel et al. (2000) is the sampling of the parameter space. We constrain the continuous five-dimensional region  $(n, \beta, M_0, r_0, P_I)$ , whereas van der Marel et al. (2000) sample discrete points within that space. They find that an NFW profile with  $n = 1$  matches the CNOC data, but they do not consider  $n = 2$  models; we find that our clusters and groups have steeper matter distributions, with the best fit model closer to  $n = 2$ .

The fixed functional form of our DF can also explain the large transition radii. Careful examination of equation (19) shows that increasing  $\gamma$ , the slope of the galaxy density as  $r \rightarrow \infty$ , also steepens the slope of the galaxy density  $n_g$  near  $r = 0$ . It is still the case that  $n_g = 2\beta$  at  $r = 0$ , but the derivative of the slope  $dn_g/dr$  becomes larger and larger as  $\gamma$  is increased. Our data have a steep galaxy density slope near  $r = 0$ , and there are only two ways to fit this: increasing  $\beta$  (and thus generating more radially anisotropic orbits) or increasing  $\gamma$  (and thus steepening the overall galaxy energy distribution and hence the galaxy density). This is where the velocity distribution enters. As Figure 6 shows, the grand total velocity histogram is not strongly peaked at  $\tilde{v}_z = 0$  (in fact, it is consistent with a Gaussian distribution). Dramatically increasing  $\beta$  would yield a velocity distribution that is much more sharply peaked than the data. Thus the only choice left the model is to increase  $\gamma$  to larger values,  $\gtrsim 8$ . However, the slope of the galaxy density is not as steep as 8 at the limits of the survey. Hence, a large transition radius  $r_0$  is required to fit the outer regions of the galaxy number density profile. At the same time, the velocity data allow the matter profile to be nearly a power law within the region constrained by the data. Note that the upper limits on  $\tilde{r}_{5/2}$  are possible because if  $\tilde{r}_{5/2}$  becomes too large, then the galaxy density profile begins to resemble a single power law as well, and the position data do not favor this limit (Figure 6).

Another shortcoming is that our method constrains the total matter density, rather than the dark matter density alone. While the dark component is thought to dominate the mass, it does not do so overwhelmingly—in typical clusters,  $\sim 25\%$  of the mass is in baryons, chiefly in the form of the X-ray emitting medium. In the innermost regions of clusters with a dominant central galaxy, it is actually the stellar mass density and not the dark mat-

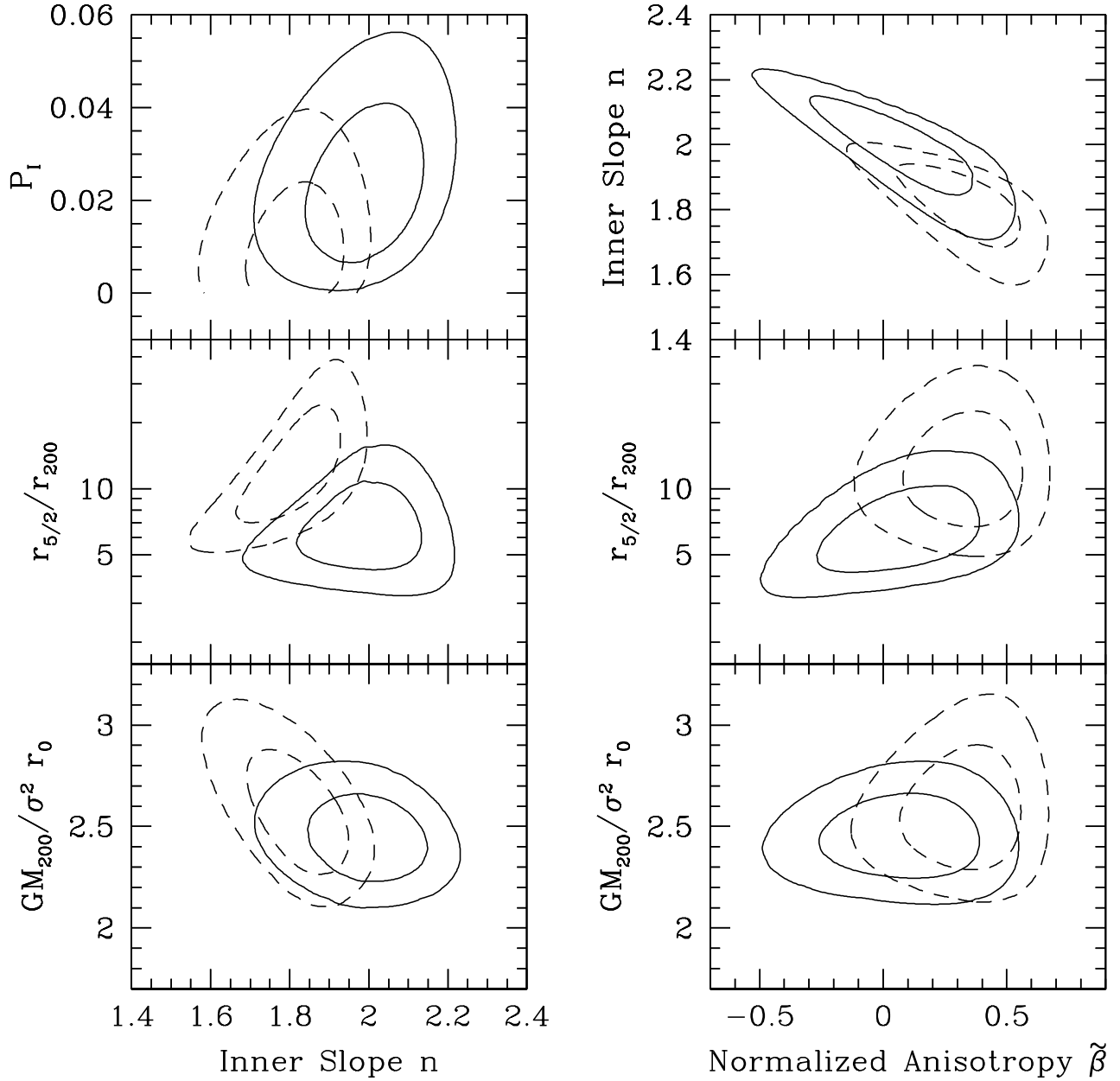


FIG. 7.— Results of the likelihood maximization for free  $n$ . Solid contours represent the low- $\sigma$  ensemble and dashed contours represent the high- $\sigma$  ensemble. Shown are marginalized joint 68% and 95% probability contours in the matter density slope  $n$ , normalized anisotropy  $\tilde{\beta}$ , mass normalization  $M_{200}$ , transition radius  $r_{5/2}$ , and interloper fraction  $P_I$ .



ter density that dominates the total distribution (Koopmans & Treu 2003). Lensing and stellar dynamical models which take into account the starlight separately from the dark component show conflicting results—while the central regions of some clusters exhibit flat  $n = 0$  (Sand et al. 2002) dark matter densities, others show evidence for steep  $n \sim 2$  dark matter profiles (Davis et al. 2003).

It would be of great value to constrain the inner density slope  $n$  independently of the slope at infinity. Unfortunately, because the data are sparse, we cannot derive true constraints on the outer galaxy density slope  $\gamma$ , and measuring the outer total matter density would be even more difficult. With  $10^5$  redshifts a true, simultaneous constraint on the outer and inner slopes would be possible; such large data sets could allow a direct measurement of the phase space distribution function (Merritt 1993), without requiring us to parameterize the shape of the distribution function, as in equation (12). Data sets that are only factors of two larger are unlikely to make that difference. Also, leveraging independent constraints on the inner and outer slopes by extending the survey beyond  $r_{200}$  is unlikely to provide better constraints. In these outer regions, most of the galaxies are experiencing spherical infall, which makes equilibrium models clearly unapplicable, and caustic modeling techniques preferable (van Haarlem et al. 1993; Diaferio 1999; Geller et al. 1999; Rines et al. 2000). However, one can envision a technique that combines equilibrium dynamics at small scales with the study of the infall caustics outside the virialized regions of the cluster. In this way one could obtain a total mass profile over several decades in  $r$ .

The most promising method of increasing the accuracy of this analysis is to extend the known membership of groups and clusters to dwarf galaxies with absolute magnitudes as faint as  $M_R = -11$ . A typical group with velocity dispersion  $\sigma \approx 300 \text{ km s}^{-1}$  is likely to have  $\approx 60$  faint dwarf members within 1 Mpc (Trentham & Tully 2002). A large spectroscopic survey of well-selected galaxies brighter than  $m_R = 22$  in our group sample would likely yield triple the current membership count and allow for much more robust constraints.

## 6. CONCLUSION

We compare the distribution of matter in rich and poor systems of galaxies, basing our analysis on a deep catalog of galaxies in groups. The catalog consists of a statistically complete sample of 2419 newly measured redshifts in the

fields of 41 systems, as well as 979 redshifts from the literature for 8 nearby rich clusters of galaxies. Most of the groups have X-ray emission detected in the RASSCALs survey (Mahdavi et al. 2000).

We construct a low- $\sigma$  ensemble group and a high- $\sigma$  ensemble cluster from these data. The final dynamical samples have 893 members in 33 groups, and 945 members in 15 clusters, with an average velocity dispersion of  $330 \text{ km s}^{-1}$  and  $800 \text{ km s}^{-1}$  respectively.

Our study is the first to derive confidence volumes in the parameter space defined by the matter density slope  $n$ , the interloper fraction  $P_I$ , the mass normalization  $M_{200}$ , the velocity anisotropy  $\beta$ , and the characteristic radius  $r_0$  for clusters and groups of galaxies. By modeling the phase space density of the galaxies directly, we place much more stringent bounds on the mass and orbital structure of these systems than would be possible through the virial theorem or the Jeans equations. However, our results are limited by the fact that we force the galaxy distribution function (DF) to be a power law in energy and angular momentum, neglecting all other forms.

We find that neither the clusters nor the groups are consistent with a matter distribution that has a flat, constant-density core ( $n = 0$ ). Instead, within  $R < 2r_{200}$ , both the low- $\sigma$  and the high- $\sigma$  systems are consistent with a single power law with slope  $n \approx 2$ . The models require both the matter and the galaxy distribution to decline steeply outside  $2 r_{200}$ . It is conceivable that choices of DF different from ours could yield substantially different results; however, the  $\chi^2$  test cannot reject the possibility that our model is correct. Our results are consistent with cold CDM simulations that predict centrally divergent matter distributions.

We thank the anonymous referee for detailed and insightful comments that significantly improved the paper. We thank Perry Berlind, Michael Calkins, and Susan Tokarz for observing and reducing most of the data via the FAST pipeline. We have made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This research was supported by the Smithsonian Institution and by NASA through a Chandra Postdoctoral Fellowship Award issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA under contract NAS 8-39073.

## REFERENCES

- Allen, S. W. 1998, *MNRAS*, 296, 392  
 Allen, S. W., Schmidt, R. W., & Fabian, A. C. 2002, *MNRAS*, 335, 256  
 Arabadjic, J. S., Bautz, M. W., & Garmire, G. P. 2002, *ApJ*, 572, 66  
 Balberg, S., Shapiro, S. L., & Inagaki, S. 2002, *ApJ*, 568, 475  
 Binney, J. & Tremaine, S. 1987, *Galactic dynamics* (Princeton, NJ, Princeton University Press, 1987, 747 p.)  
 Biviano, A. & Girardi, M. 2003, *ApJ*, 585, 205  
 Biviano, A., Girardi, M., Giuricin, G., Mardirossian, F., & Mezzetti, M. 1993, *ApJ*, 411, L13  
 Burkert, A. 2000, *ApJ*, 534, L143  
 Carlberg, R. G., Yee, H. K. C., Ellingson, E., Abraham, R., Gravel, P., Morris, S., & Pritchet, C. J. 1996, *ApJ*, 462, 32  
 Carlberg, R. G., Yee, H. K. C., Ellingson, E., Morris, S. L., Abraham, R., Gravel, P., Pritchet, C. J., Smecker-Hane, T., Hartwick, F. D. A., Hesser, J. E., Hutchings, J. B., & Oke, J. B. 1997, *ApJ*, 485, L13  
 Cirimele, G., Nesci, R., & Trevese, D. 1997, *ApJ*, 475, 11  
 Clowe, D., Luppino, G. A., Kaiser, N., & Gioia, I. M. 2000, *ApJ*, 539, 540  
 Clowe, D. & Schneider, P. 2002, *A&A*, 395, 385  
 Colless, M., Dalton, G., Maddox, S., Sutherland, W., Norberg, P., Cole, S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Collins, C., Couch, W., Cross, N., Deeley, K., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Price, I., Seaborne, M., & Taylor, K. 2001, *MNRAS*, 328, 1039  
 Cuddeford, P. 1991, *MNRAS*, 253, 414

- Danese, L., de Zotti, G., & di Tullio, G. 1980, *A&A*, 82, 322
- Davé, R., Spergel, D. N., Steinhardt, P. J., & Wandelt, B. D. 2001, *ApJ*, 547, 574
- Davis, A. N., Huterer, D., & Krauss, L. M. 2003, *MNRAS*, 344, 1029
- Dehnen, W. 1993, *MNRAS*, 265, 250
- den Hartog, R. & Katgert, P. 1996, *MNRAS*, 279, 349
- Diaferio, A. 1999, *MNRAS*, 309, 610
- Durret, F., Adami, C., Gerbal, D., & Pislari, V. 2000, *A&A*, 356, 815
- Durret, F., Forman, W., Gerbal, D., Jones, C., & Vikhlinin, A. 1998, *A&A*, 335, 41
- Ebeling, H., Edge, A. C., Bohringer, H., Allen, S. W., Crawford, C. S., Fabian, A. C., Voges, W., & Huchra, J. P. 1998, *MNRAS*, 301, 881
- Ebeling, H., Voges, W., Bohringer, H., Edge, A. C., Huchra, J. P., & Briel, U. G. 1996, *MNRAS*, 281, 799
- Fabian, A. C., Mushotzky, R. F., Nulsen, P. E. J., & Peterson, J. R. 2001, *MNRAS*, 321, L20
- Fabricant, D., Cheimets, P., Caldwell, N., & Geary, J. 1998, *PASP*, 110, 79
- Fabricant, D. G., Kent, S. M., & Kurtz, M. J. 1989, *ApJ*, 336, 77
- Fricke, W. 1952, *Astronomische Nachrichten*, 280, 193
- Fukushige, T. & Makino, J. 1997, *ApJ*, 477, L9
- Gavazzi, R., Fort, B., Mellier, Y., Pello, R., & Dantel-Fort, M. 2003, *A&A*, 403, 11
- Geller, M. J., Diaferio, A., & Kurtz, M. J. 1999, *ApJ*, 517, L23
- Girardi, M., Escalera, E., Fadda, D., Giuricin, G., Mardirossian, F., & Mezzetti, M. 1997, *ApJ*, 482, 41
- Girardi, M., Fadda, D., Giuricin, G., Mardirossian, F., Mezzetti, M., & Biviano, A. 1996, *ApJ*, 457, 61
- Girardi, M. & Giuricin, G. 2000, *ApJ*, 540, 45
- Heisler, J., Tremaine, S., & Bahcall, J. N. 1985, *ApJ*, 298, 8
- Helsdon, S. F. & Ponman, T. J. 2000, *MNRAS*, 315, 356
- Hernquist, L. 1990, *ApJ*, 356, 359
- Hickson, P. 1982, *ApJ*, 255, 382
- Jaffe, W. 1983, *MNRAS*, 202, 995
- Jing, Y. P. & Suto, Y. 2000, *ApJ*, 529, L69
- Kent, S. M. & Gunn, J. E. 1982, *AJ*, 87, 945
- Kolokotronis, V., Basilakos, S., Plionis, M., & Georgantopoulos, I. 2001, *MNRAS*, 320, 49
- Koopmans, L. V. E. & Treu, T. 2003, *ApJ*, 583, 606
- Kurtz, M. J. & Mink, D. J. 1998, *PASP*, 110, 934
- Kurtz, M. J., Mink, D. J., Wyatt, W. F., Fabricant, D. G., Torres, G., Kriss, G. A., & Tonry, J. L. 1992, in *ASP Conf. Ser.* 25: *Astronomical Data Analysis Software and Systems I*, Vol. 1, 432
- Lewis, A. D., Buote, D. A., & Stocke, J. T. 2003, *ApJ*, 586, 135
- Lewis, A. D., Ellingson, E., Morris, S. L., & Carlberg, R. G. 1999, *ApJ*, 517, 587
- Lupton, R. 1993, *Statistics in theory and practice* (Princeton, N.J.: Princeton University Press, —c1993)
- Mahdavi, A. 2001, *ApJ*, 546, 812
- Mahdavi, A., Böhringer, H., Geller, M. J., & Ramella, M. 2000, *ApJ*, 534, 114
- Mahdavi, A., Geller, M. J., Böhringer, H., Kurtz, M. J., & Ramella, M. 1999, *ApJ*, 518, 69
- Markevitch, M., Vikhlinin, A., Forman, W. R., & Sarazin, C. L. 1999, *ApJ*, 527, 545
- Merritt, D. 1985, *AJ*, 90, 1027
- , 1993, *ApJ*, 413, 79
- Merritt, D. & Saha, P. 1993, *ApJ*, 409, 75
- Metzler, C. A., White, M., & Loken, C. 2001, *ApJ*, 547, 560
- Moore, B., Quinn, T., Governato, F., Stadel, J., & Lake, G. 1999, *MNRAS*, 310, 1147
- Nakano, T. & Makino, J. 1999, *ApJ*, 510, 155
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *NFW*, *ApJ*, 490, 493
- Neill, J. D., Brodie, J. P., Craig, W. W., Hailey, C. J., & Misch, A. A. 2001, *ApJ*, 548, 550
- Neumann, D. M. & Böhringer, H. 1997, *MNRAS*, 289, 123
- Oegerle, W. R., Hill, J. M., & Fitchett, M. J. 1995, *AJ*, 110, 32
- Pisani, A. 1993, *MNRAS*, 265, 706
- Ponman, T. J., Cannon, D. B., & Navarro, J. F. 1999, *Nature*, 397, 135
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical recipes in FORTRAN. The art of scientific computing* (Cambridge: University Press, —c1992, 2nd ed.)
- Rines, K., Geller, M. J., Diaferio, A., Mohr, J. J., & Wegner, G. A. 2000, *AJ*, 120, 2338
- Rines, K., Geller, M. J., Kurtz, M. J., & Diaferio, A. 2003, *AJ*, 126, 2152
- Rines, K., Mahdavi, A., Geller, M. J., Diaferio, A., Mohr, J. J., & Wegner, G. 2001, *ApJ*, 555, 558
- Sand, D. J., Treu, T., & Ellis, R. S. 2002, in 5 pages, 3 figures, accepted for publication in *ApJ. Letters*, 7048+
- Sheldon, E. S., Annis, J., Böhringer, H., Fischer, P., Frieman, J. A., Joffe, M., Johnston, D., McKay, T. A., Miller, C., Nichol, R. C., Stebbins, A., Voges, W., Anderson, S. F., Bahcall, N. A., Brinkmann, J., Brunner, R., Csabai, I., Fukugita, M., Hennessy, G. S., Ivezić, Ž., Lupton, R. H., Munn, J. A., Pier, J. R., & York, D. G. 2001, *ApJ*, 554, 881
- Slezak, E., Durret, F., & Gerbal, D. 1994, *AJ*, 108, 1996
- Soker, N., White, R. E., David, L. P., & McNamara, B. R. 2001, *ApJ*, 549, 832
- Tonry, J. & Davis, M. 1979, *AJ*, 84, 1511
- Tremaine, S., Richstone, D. O., Byun, Y., Dressler, A., Faber, S. M., Grillmair, C., Kormendy, J., & Lauer, T. R. 1994, *AJ*, 107, 634
- Trentham, N. & Tully, R. B. 2002, *MNRAS*, 335, 712
- Tyson, J. A., Kochanski, G. P., & dell'Antonio, I. P. 1998, *ApJ*, 498, L107
- van der Marel, R. P., Magorrian, J., Carlberg, R. G., Yee, H. K. C., & Ellingson, E. 2000, *AJ*, 119, 2038
- van Haarlem, M. P., Cayon, L., Guiterrez de La Cruz, C., Martinez-Gonzalez, E., & Rebolo, R. 1993, *MNRAS*, 264, 71
- Wilkinson, M. I., Kleyna, J., Evans, N. W., & Gilmore, G. 2002, *MNRAS*, 330, 778
- Yee, H. K. C., Ellingson, E., & Carlberg, R. G. 1996, *ApJS*, 102, 269
- Zabludoff, A. I., Huchra, J. P., & Geller, M. J. 1990, *ApJS*, 74, 1

TABLE 1  
NEWLY OBSERVED SYSTEMS

Name	$\alpha_{2000}$	$\delta_{2000}$	$N_{\text{obs}}$	$N_{\text{memb}}$	$c\bar{z}$	$\sigma_z$	$\log L_X$	Other ID
SRGb062 <sup>a</sup>	00:18:25.2	+30:04:13.8	84	45	6818±54	379±43	42.66±0.14	
SRGb063	00:21:38.4	+22:24:20.5	75	45	5771±48	336±36	42.56±0.18	
SRGb090 <sup>b</sup>	01:07:02.2	+32:19:21.7	142	104	5150±47	508±34	42.81±0.08	NGC 383 Group
SRGb102 <sup>b</sup>	01:25:28.8	+01:45:17.0	109	35	5433±66	409±41	42.56±0.12	NGC 533 Group
SRGb119 <sup>a</sup>	01:56:21.6	+05:37:04.7	54	29	5484±64	372±42	42.30±0.20	NGC 741 Group
SRGb145	02:31:48.0	+01:16:27.2	55	37	6714±61	387±43	42.34±0.31	
SRGb149	02:38:43.8	+02:01:11.4	68	44	6670±54	376±38	< 41.98	
SRGb155 <sup>b</sup>	02:52:48.7	-01:17:02.5	66	26	7259±123	650±93	42.69±0.20	
SRGb158	02:55:51.0	+09:18:48.4	33	22	7810±61	301±36	< 41.98	
NRGb004 <sup>a</sup>	08:38:11.5	+25:07:00.1	37	20	8478±91	440±144	42.937±0.14	
NRGb007 <sup>a</sup>	08:49:31.9	+36:35:40.1	28	11	7625±94	322±81	< 41.979	
NRGb025 <sup>a</sup>	09:13:28.3	+30:06:13.7	35	31	6735±81	458±65	< 41.86	
NRGs027	09:16:06.2	+17:36:08.3	59	36	8615±63	387±54	42.62±0.18	
NRGb032 <sup>a</sup>	09:19:48.0	+33:45:32.8	67	45	6834±64	459±63	42.69±0.14	Abell 779
NRGs038 <sup>b</sup>	09:23:35.0	+22:19:58.4	65	56	9218±104	784±60	42.47±0.24	
NRGb043 <sup>a</sup>	09:28:09.9	+30:03:35.0	29	25	7884±47	257±32	< 42.07	
NRGb045 <sup>a</sup>	09:33:27.1	+34:03:02.9	26	9	8211±36	110±31	42.11±0.42	
NRGb057 <sup>a</sup>	09:44:28.1	+36:08:55.7	24	13	6766±45	168±25	< 41.76	
SS2b144	09:49:59.9	-05:02:48.3	35	21	6303±43	202±22	< 41.88	
NRGs076 <sup>b</sup>	10:06:41.8	+14:25:49.8	60	40	9273±150	929±145	42.34±0.30	
NRGb078	10:13:53.5	+38:44:24.4	87	39	6770±51	335±35	42.16±0.52	
NRGs110 <sup>b</sup>	11:00:50.9	+10:33:17.3	60	38	10579±76	493±62	42.75±0.20	Abell 1142
NRGs117 <sup>b</sup>	11:10:31.4	+28:43:39.4	86	80	9814±73	667±47	42.98±0.13	Abell 1185
NRGs127 <sup>a</sup>	11:21:30.5	+34:13:21.5	12	10	10485±64	206±36	< 42.13	
SS2b164	11:22:43.0	-07:44:54.2	88	47	7107±51	361±46	42.33±0.27	
NRGs156 <sup>a</sup>	11:46:16.3	+33:11:03.1	35	28	9640±58	320±37	< 41.97	
NRGb177	12:04:17.8	+20:15:18.0	113	74	6995±48	416±35	42.52±0.16	NGC 4065 Group
NRGb181 <sup>a</sup>	12:07:07.4	+31:19:46.5	17	13	7053±118	426±70	< 41.78	
NRGb184	12:08:01.0	+25:15:13.7	113	52	6726±48	376±40	42.30±0.24	
NRGs241 <sup>b</sup>	13:20:16.6	+33:08:13.2	59	40	10980±78	503±45	42.92±0.12	
NRGb244 <sup>a</sup>	13:24:10.8	+13:58:47.6	33	19	6947±56	268±33	42.63±0.18	NGC 5129 Group
NRGb247	13:29:31.2	+11:47:19.0	58	39	6850±61	395±43	42.67±0.18	NGC 5171 Group
NRGb251 <sup>a</sup>	13:34:25.7	+34:40:54.8	48	26	7345±50	265±30	42.44±0.17	
SS2b239	13:49:10.8	-07:18:11.7	69	24	7329±59	311±29	42.38±0.26	HCG 67
NRGb302	14:28:33.1	+11:22:07.7	53	31	7892±58	324±32	42.05±0.59	
NRGs317 <sup>a</sup>	14:47:09.8	+13:42:23.4	39	19	8898±72	338±38	42.51±0.30	
SRGb009 <sup>a</sup>	22:14:48.0	+13:50:17.5	46	35	7775±57	349±33	42.44±0.26	
SRGb013 <sup>b</sup>	22:50:00.7	+11:40:15.6	55	30	7689±101	578±76	42.45±0.21	
SRGb016	22:58:14.2	+25:56:13.0	58	43	7394±46	327±33	< 41.84	
SRGb037	23:28:46.6	+03:30:49.1	85	20	5221±86	413±59	41.91±0.44	
SS2b312	23:47:24.0	-02:19:08.4	59	27	6773±47	274±42	42.41±0.20	HCG 97

Note. —  $N_{\text{obs}}$  is the number of galaxies whose redshifts were measured in this work;  $N_{\text{memb}}$  is the number we identify as group members;  $\bar{z}$  is the mean group redshift, and  $\sigma_z$  is the velocity dispersion in  $\text{km s}^{-1}$ , with errors from standard bootstrap analysis.  $L_X$  is the 0.1-2.4 keV luminosity in  $h_{100}^{-2} \text{ erg s}^{-1}$  as calculated in Mahdavi et al. (2000).

<sup>a</sup>The data for these groups appeared originally in Mahdavi et al. (1999); however, our algorithm for determining group membership differs from the one used in that work, and hence the total number of members varies slightly. Also,  $N_{\text{obs}}$  in our table differs from  $N_{\text{Total}}$  in Mahdavi et al. (1999), because we list the actual number of galaxies observed, rather than the number in the field brighter than a given magnitude.

<sup>b</sup>Because of their large velocity dispersions, these systems were grouped with the rich clusters in Table 3 for our analysis.

TABLE 2  
NEW GALAXY POSITIONS AND VELOCITIES <sup>a</sup>

Name	$\alpha_{2000}$			$\delta_{2000}$			$cz$	$\epsilon_{cz}$	$d$	Member
SRGb062.001	00	11	57.0	+29	29	08	28038	35	1.808	N
SRGb062.002	00	12	11.8	+29	19	09	7728	21	1.838	N
SRGb062.003	00	12	17.5	+29	52	17	6921	20	1.595	N
SRGb062.004	00	12	28.4	+29	32	38	6862	15	1.656	N
SRGb062.005	00	12	38.3	+30	06	08	6791	23	1.487	M
SRGb062.006	00	12	45.0	+29	22	15	10419	15	1.683	N
SRGb062.007	00	13	12.7	+31	08	43	14465	28	1.846	N
SRGb062.008	00	13	45.1	+30	11	40	7109	28	1.209	M

Note. —  $cz$  is the galaxy redshift times the speed of light in  $\text{km s}^{-1}$ ,  $\epsilon_{cz}$  is the measurement error, and  $d$  is the distance from the X-ray center in Mpc assuming  $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . Members are indicated by an “M” in the Member column; nonmembers are indicated by an “N.”

<sup>a</sup>This table is for illustrative purposes only. The complete data appear in the electronic version of the Journal.

TABLE 3  
CLUSTER SAMPLE

Name	$\alpha_{2000}$	$\delta_{2000}$	$N_{\text{memb}}$	$c\bar{z}$	$\sigma_z$	$\log L_X$	References	
Abell 262	01:52:50.4	36:08:46	103	4920±55	560±41	43.14	0.58	G97,N01
Abell 3158	03:42:39.6	-53:37:50	118	17740±91	964±56	44.12	1.70	K01,G96
Abell 496	04:33:37.1	-13:14:46	227	9920±53	729±37	43.95	1.17	D00,M99
Abell 1644	12:57:14.8	-17:21:13	85	14210±100	947±73	43.94	...	G97
Abell 1795	13:49:00.5	26:35:07	84	18880±95	879±77	44.44	1.76	G97
Abell 1809	13:53:18.9	05:09:15	61	23750±100	747±67	43.60	...	G97
Abell 3571	13:47:28.9	-32:51:57	90	11700±101	969±69	44.26	1.72	G96
Abell 2029	15:10:58.7	05:45:42	87	23140±140	1235±78	44.58	2.20	O95,S94

Note. —  $N_{\text{memb}}$  is the number we identify as cluster members;  $\bar{z}$  is the mean cluster redshift, and  $\sigma_z$  is the velocity dispersion of our derived membership in  $\text{km s}^{-1}$ , with errors from standard bootstrap analysis.  $L_X$  is the total 0.1-2.4 keV luminosity in  $h_{100}^{-2} \text{ erg s}^{-1}$  from Ebeling et al. (1996), the typical error is 20%. The references are to X-ray and optical studies of the clusters that suggest they could be dynamically relaxed. C97: Cirimele et al. (1997); D00: Durret et al. (2000); G96: Girardi et al. (1996); G97: Girardi et al. (1997); K01: Kolokotronis et al. (2001); O95: Oegerle et al. (1995); N01: Neill et al. (2001); S94: Slezak et al. (1994).

TABLE 4  
SIMULATIONS OF THE MEASUREMENT TECHNIQUE

ID	$\tilde{\beta}$			$n$		$\tilde{r}_{5/2}$		$\tilde{M}_{200}$		$P_I$		Fit Quality
	In	Out		In	Out	In	Out	In	Out	In	Out	
A	0.0	-0.3	- 0.3	1.0	0.6–1.3	1.0	0.8–1.2	0.3	0.27–0.35	0.05	0.04–0.07	0.96
B	0.5	-0.2	- 0.7	2.0	1.9–2.3	1.5	1.2–1.8	3	1.3–5.0	0.1	0.08–0.15	0.75
C	-0.5	-1.0	- -0.3	0.0	0.0–1.3	2.0	1.8–2.2	1	0.6–2.0	0.01	0.0–0.03	0.20
D	0.0	-0.5	- 0.2	0.0	0.0–0.7	1.0	0.7–1.0	0.4	0.3–0.4	0.03	0.01–0.04	0.34

Note. — Simulations of 893 galaxies carried out with  $\gamma = 8$ . “In” refers to the input parameter set; “Out” shows the recovered 1D 95% marginalized confidence interval for each parameter.

TABLE 5  
SUMMARY OF RESULTS

	Low- $\sigma$	High- $\sigma$
<i>General Statistics</i>		
No. of systems	33	15
Completeness limit	$1.8r_{200}$	$0.8r_{200}$
No. of members within $0.8r_{200}$	521	945
No. of members within $1.8r_{200}$	893	...
Range in $\sigma_z$ (km/s)	110-470	470-1240
Gaussian fit to $\tilde{v}_z$	$\chi^2/\nu = 22/27$	$\chi^2/\nu = 32/28$
KS Test in $R^a$	$d = 0.106, p = 10^{-3}$	
KS Test in $\tilde{v}_z$	$d = 0.028, p = 0.87$	
<i>Likelihood of Constant Density Core (<math>n = 0</math>)</i>		
Fit Quality, $\beta \leq 0$	$q = 0.003$	$q < 10^{-5}$
Fit Quality, $\beta$ free	$q = 0.02$	$q = 0.003$
<i>Likelihood of Free <math>n</math></i>		
Minimum $\gamma$	8	12
Fit Quality	$q = 0.169$	$q = 0.124$
$n$	1.8-2.2	1.6-2.0
$\tilde{\beta}$	-0.37 - 0.42	0.02 - 0.67
$r_{5/2}/r_{200}$	3-12	5.8-26
$GM_{200}/\sigma^2 r_0$	2.2-2.7	2.2-3.0
$P_I$	0.01-0.05	0.01-0.02

<sup>a</sup>Conducted only for galaxies with  $R < 0.8r_{200}$ .

## APPENDIX

## THE DISTRIBUTION FUNCTION INTEGRAL

In the statistical interpretation of the distribution function (DF), the probability of observing a galaxy with a line-of-sight peculiar velocity  $v_z$  at a projected distance  $R$  from the cluster center in a survey of  $N_0$  galaxies within a limiting radius  $R_{\text{lim}}$  is (Merritt & Saha 1993; van der Marel et al. 2000)

$$p'(R, v_z) = \frac{2\pi R}{N_0} \int \int \int f(\mathcal{E}, L^2) dv_\phi dv_R dz, \quad (\text{A1})$$

where  $(v_R, v_\phi, v_z)$  are the components of the galaxy's peculiar motion in cylindrical coordinates,  $\mathcal{E} \equiv \Psi(r) - v^2/2$ ,  $L$  is the angular momentum, and  $f$  is the distribution function (DF). The probability distribution is normalized such that

$$\int_0^{R_{\text{lim}}} \int_{-\infty}^{\infty} p'(R, v_z | \mathbf{a}) dv_z dR = \int_0^{R_{\text{lim}}} \frac{2\pi R}{N_0} \Sigma(R) dR = 1, \quad (\text{A2})$$

where  $\Sigma(R)$  is the surface number density of galaxies. We use a DF of the form

$$f(\mathcal{E}, L^2) = f_0 \mathcal{E}^{\alpha-1/2} L^{-2\beta} \quad (\text{A3})$$

where  $f_0$  is fixed by the normalization requirement in equation (A2). The limits of the integration are determined as follows: (1) a galaxy bound to the cluster must always have  $v_R^2 + v_\phi^2 + v_z^2 \leq 2\Psi$ , and (2) a galaxy with peculiar velocity  $v_z$  has an associated maximum radius  $r_{\text{max}}$  such that  $v_z^2 = 2\Psi(r_{\text{max}})$ , beyond which it is unbound. With these conditions, and switching to the coordinate system  $v_R = w \sin \eta$ ,  $v_\phi = w \cos \eta$ , we have

$$p'(R, v_z) = \frac{2\pi R f_0}{N_0} \int_R^{r_{\text{max}}} \frac{2r dr}{\sqrt{r^2 - R^2}} \int_0^{2\pi} d\eta \times \quad (\text{A4})$$

$$\int_0^{\sqrt{2\Psi - v_z^2}} [\Psi(r) - w^2/2 - v_z^2/2]^{\alpha-1/2} L^{-2\beta} w dw,$$

$$L^2 = r^2 w^2 \cos^2 \eta + r^2 (w \sin \eta \cos \theta - v_z \sin \theta)^2, \quad (\text{A5})$$

$$\sin \theta \equiv \frac{R}{r}. \quad (\text{A6})$$

For  $\beta > 0$ , the above expression has an integrable singularity at  $L = 0$ , or  $(\eta = \pi/2, w = v_z \tan \theta)$ . Changing variables from  $(\eta, w)$  to  $(Q, L^2)$  is tempting, but the integration boundaries in those coordinates are not necessarily analytic. It is best to work in the  $(\eta, w)$  plane and take advantage of the fact that, for Gaussian quadrature, the innermost integrand will never be evaluated at the singular point  $\eta = \pi/2$ , and so will always be finite at  $w = v_z \tan \theta$ . We find that making the change of variables appropriate for an inverse square-root singularity in  $w$  ( $\omega = w^2$ ) (Press et al. 1992) yields accurate results for the innermost integral. As for the middle,  $\eta$  integral, it may be evaluated using Gaussian quadrature (Press et al. 1992) with a constant weighting function for  $\beta < 1/2$ . For  $\beta \geq 1/2$ , Gaussian quadrature, with  $(\pi/2 - \eta)^{1-2\beta}$  as the weighting function, is best.

The galaxy number density  $\nu(r)$  corresponding to the chosen DF is given by an integral over all velocities; this integral is discussed in detail by Cuddeford (1991) and Merritt (1985).

$$\nu(r) = \int_{v^2 \leq 2\Psi} f(\mathcal{E}, L^2) d^3\mathbf{v} \quad (\text{A7})$$

$$= \frac{2\pi f_0}{r} \int_0^\Psi \mathcal{E}^{\alpha-1/2} d\mathcal{E} \int_0^{2r^2(\Psi-\mathcal{E})} \frac{L^{-2\beta} dL^2}{\sqrt{2r^2(\Psi-\mathcal{E}) - L^2}} \quad (\text{A8})$$

$$= 2^{3/2-\beta} \pi^{3/2} f_0 \frac{\Gamma(\alpha+1/2)\Gamma(1-\beta)}{\Gamma(\alpha-\beta+2)} r^{-2\beta} \Psi(r)^{\alpha-\beta+1}. \quad (\text{A9})$$